Substrate Recognition by Protein-only Ribonuclease P

by

Bradley Phillip Klemm

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Biological Chemistry)
in the University of Michigan
2017

Doctoral Committee:

Professor Carol A. Fierke, Chair
Assistant Professor Peter L. Freddolino
Associate Professor Aaron C. Goldstrohm, University of Minnesota
Associate Professor Patrick J. O'Brien
Professor Nils G. Walter

DEDICATION

To Danielle, who supported me through the long hours writing and late nights in lab that resulted in this work.

ACKNOWLEDGEMENTS

I would like to begin by thanking all of those who mentored me over the years. My advisor Carol Fierke has been an excellent mentor and I know that my scientific training benefited from her guidance. All members of my thesis committee, past and present, were invaluable for their advice and guidance. I thank David Engelke for serving on my committee for 2 years and for numerous helpful discussions about my work, for which Chapter 3 is greatly improved. I benefited from wonderful instructors during my time as an undergraduate at UW – Eau Claire, all of whom were fundamental in laying the foundations of my scientific career. David Lewis – I affirmed my interest in science and developed my laboratory skills while working with you. While I may have turned away from chemistry and to the dark side, it was your advice that brought me to Michigan and for that I'm grateful.

I would like to thank all the members of the Fierke lab RNase P sub-group for the discussions we've shared. Michael Howard and I spent many hours discussing science, life, PRORPs and experiments. Yu Chen, Xin Liu, Nancy Wu, and Kipchumba Kaitany were always available for critical discussions and supplying various materials. I would also like to thank Xin Liu for deriving an equation that helped us address reviewer comments for Chapter 2 and strengthen the data in Appendix A. I also want to thank Matthew Henley for bringing to my attention the Staudinger reaction and thus solving a months-long issue with proteins containing a non-natural amino acid substitution.

I would like to extend thanks to all other members of the Fierke lab who have helped me over the years with their valuable advice and assistance with experiments. My discussions with Elia Wright and Eric Sullivan about life and the minutia of enzyme kinetics were a great help. Andrea Stoddard regularly assisted

me by lending her expertise in molecular biology. I'm certain I would not have accomplished nearly as much in the lab if it weren't for her guidance.

I'm grateful for all the time I've spent with my family and friends, new and old. I apologize to my old family and friends for my 5 year absence, but appreciate your understanding. I've benefited greatly from a host of wonderful new friends in Ann Arbor. You all made the transition to graduate school much easier. Every Saturday tailgating for Michigan football (Go Blue!) and every tubing trip down the Huron came as a welcome break. To my new family, I'm looking forward to many more years with you all.

I'm grateful to my parents for your love and support over the years. You guided me through my formative years and your values helped me to continue in the face of the repeated set-backs that are typical of all scientific endeavors. I'm glad for my new parents; I now have two wonderful sets of parents and appreciate you all.

Last, but surely not least, I would like to thank my wife Danielle. I appreciate the support you've given me, even through your own thesis. Despite completing two theses, we've managed to explore new states, Ontario, and even get married. Because of you, this degree is only the second best thing that came from my time in Michigan.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SCHEMES

LIST OF APPENDECIES

ABSTRACT

Ribonuclease P (RNase P) is a metal-dependent endonuclease responsible for the 5’ end maturation of precursor transfer ribonucleic acids (pre-tRNAs). All domains of life encode an RNA-dependent RNase P, with a conserved RNA component that is fully catalytic in Bacteria. Many Eukarya also encode protein-only RNase Ps (PRORPs), which include a unique class of nuclease domains. The PRORPs in metazoans, such as humans, process the pre-tRNAs encoded by the mitochondrial genome. These PRORPs include three subunits: the nuclease, a tRNA methyltransferase, and a dehydrogenase/reductase. In contrast, the PRORP nucleases found in trypanosomes, plants, and algae do not require additional subunits for catalysis. The PRORPs from the land plant *Arabidopsis thaliana* are the only RNase Ps responsible for 5’ end maturation of pre-tRNAs in all three tRNA-encoding cellular compartments. In this thesis, we investigate the strategies employed by plant PRORPs to recognize their substrates.

In Chapter 2, we define the structural features of pre-tRNAs that are required for PRORP recognition. We find that the anti-codon stem is dispensable for activity, while the D-loop is required. Furthermore, the 5’ leader sequence beyond two nucleotides and the entire 3’ trailer sequence do not increase substrate affinity or significantly affect catalysis. The three differentially localized PRORPs encoded by *Arabidopsis* do not have specificity for substrates based on compartment of origin. We also identify an alternative 5’ end selection behavior by PRORPs *in vitro*. Plant PRORPs recognize an acceptor stem with an extra base pair and miscleave these substrates, but subsequently re-process them to the correct 5’ end *in vitro*.

In Chapter 3, we investigate the molecular determinants of the recognition complex of pre-tRNA substrates and plant PRORPs. Increasing ionic strength

using NaCl reduces the binding affinity and observed rate constants for *Arabidopsis* PRORPs. These dependencies reveal at least 4–5 direct contacts to backbone phosphates. Applying the same experiments to short-leader substrates reveals 1 phosphate contact in the 5’ leader. The catalytic inhibition by NaCl suggests that the PRORP kinetic mechanism is more complicated than previously proposed. We also identify a novel tRNA-recognition surface on the PRORP pentatricopeptide repeat (PPR) domain that is conserved among PRORPs. The residues that contact substrate are different from those used by single-stranded RNA binding PPR proteins.

In Chapter 4, we develop a method to map the PRORP-substrate recognition complex. We use photo-activatable non-natural amino acid (NNAA) cross-linkers to yield a covalent PRORP-substrate complex. We observe an increase in apparent molecular weight specific for the NNAA-substituted PRORPs and consistent with tRNA-sized molecules when photo-activated within *E*. *coli*. We use a primer extension assay to identify the cross-linked sites on the pre-tRNA substrate. The method we develop could be adapted for use with RNA sequencing methods to identify *in vivo* substrates for PRORPs or other RNA-binding proteins.

Overall, this work provides key insights into substrate recognition by an important class of endonucleases, including molecular determinants of substrate recognition, and develops a method to expand and continue these studies.

CHAPTER 1

Introduction†

ABSTRACT

Ribonuclease P (RNase P) is an essential endonuclease responsible for catalyzing 5' end maturation in precursor transfer RNAs. Since its discovery in the 1970s, RNase P enzymes have been identified and studied throughout the three domains of life. Interestingly, RNase P is either RNA-based, with a catalytic RNA subunit, or a protein-only (PRORP) enzyme with differential evolutionary distribution. The available structural data, including the active site data, provide insight into catalysis and substrate recognition. The hydrolytic and kinetic mechanisms of the two forms of RNase P enzymes are similar, yet features unique to the RNA-based and PRORP enzymes are consistent with different evolutionary origins. The various RNase P enzymes, in addition to their primary role in tRNA 5' maturation, catalyze cleavage of a variety of alternative substrates, indicating a diversification of RNase P function *in vivo*.

INTRODUCTION

The 1989 Nobel Prize in Chemistry was awarded to Sidney Altman and Thomas Cech "for their discovery of catalytic properties of RNA" (1). Altman was recognized for determining that the *Escherichia coli* ribonuclease P (RNase P) activity originates from an RNA subunit (2, 3). RNase P is a metal-dependent

---

†This chapter is adapted from reference: Klemm, BP; Wu, N; Chen, Y; Liu, X; Kaitany, KJ; Howard, MJ; Fierke, CA. The Diversity of Ribonuclease P: Protein and RNA Catalysts with Analogous Biological Functions. *Biomolecules* 2016, 6(2):E27.

B.P.K. organized the review. B.P.K., N.W., Y.C., X.L., and K.J.K. wrote the paper. M.J.H. performed the modeling experiments for Fig. 1-7. All authors edited the paper.

endonuclease that catalyzes phosphodiester bond hydrolysis of precursor transfer RNA (pre-tRNA), generating tRNA with a mature 5' end, including a 5' phosphate, and a 5' leader with a 3' hydroxyl (Fig. 1-1) (4). Pre-tRNAs are non-functional in translation and must be processed for protein synthesis to occur (5, 6). Therefore, RNase P plays a key role in cellular homeostasis and survival.

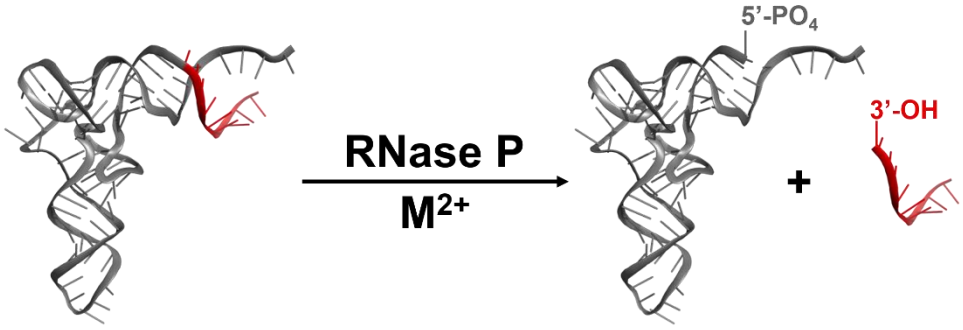


Figure 1-1: RNase P enzymes catalyze metal-dependent, endonucleolytic cleavage of pre-tRNA (adapted with permission from (7)).

Two types of RNase P enzymes exist: 1) RNA-dependent enzymes (ribozymes), for which the active site is located in a catalytic RNA subunit, and 2) protein-only RNase Ps (PRORPs; also referred to as "proteinaceous RNase Ps" in various literature sources, we note that eukaryotic RNA-based RNase Ps are also proteinaceous, with ≥ 70% protein by mass, so "protein-only RNase P" is more precise and we favor this nomenclature). RNase Ps provide the opportunity to compare catalytic strategies of independently evolved protein and RNA catalysts in the only known biological model system in which both are utilized in extant biology to execute the same biological function.

In the following sections, we will review: the diversity and distribution of RNase Ps throughout life; the known structures of RNase P enzymes; the catalytic strategies employed by RNase P; and the substrate recognition strategies of RNase P.

DIVERSITY AND DISTRIBUTION OF RNase Ps

RNase P enzymes are found in all domains of life and in nearly all species. There is one known exception: the obligate symbiont *Nanoarchaeum equitans* does not encode pre-tRNAs with 5' extensions and thus lacks an RNase P (8, 9).

While RNA-based RNase Ps have been found in all three domains of life, PRORPs are found exclusively in eukaryotes (Fig. 1-2) (10). Given the broad distribution of PRORPs across Eukarya, in particular across lineages as diverse as trypanosomes, plants, and animals, it is likely that the proteins emerged in basal eukaryotes and prior to the divergence of supergroups. However, there are several eukaryotic lineages for which an RNase P enzyme has not been identified in the genome–*Jacobida* (Excavata), Cyanidiophyceae (Archaeplastida), and *Nuclearia* (Opisthokonta)–highlighted by an apparent lack of either PRORP or RNase P RNA sequences (11). These species may have maintained a P RNA that is too distinct from the canonical sequence to be identified from genomic queries, or have genomic sequences that are currently too incomplete to allow identification of a P RNA or PRORP.

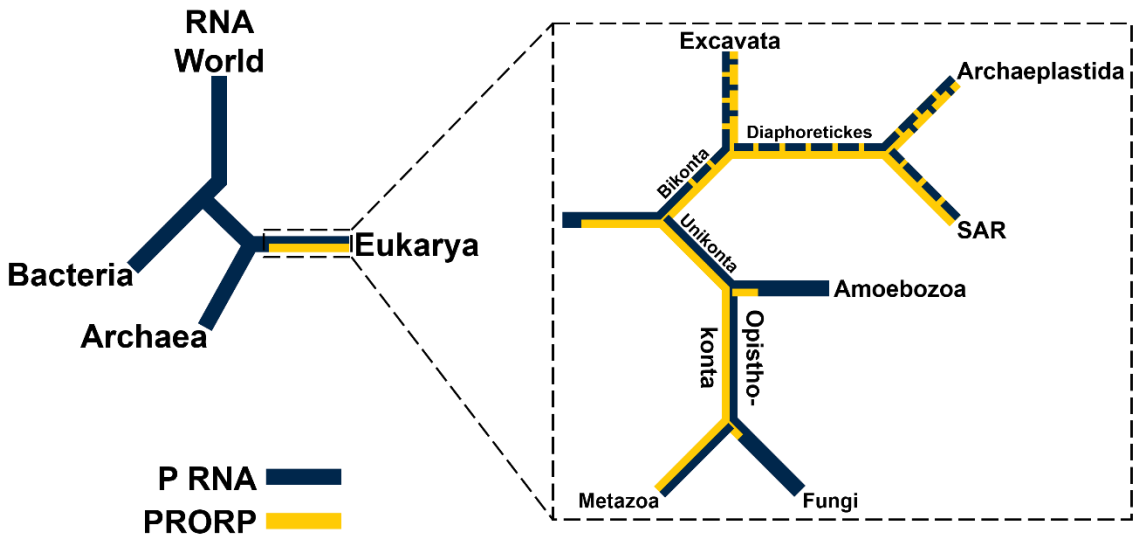


Figure 1-2: Distribution of RNase P enzymes across evolutionary lineages. Presence of a P RNA (blue) or PRORP (yellow) is indicated. (**Left**) RNA-based RNase Ps are found in all domains of life and are proposed to have evolved from an RNA predecessor lacking protein components. PRORPs exist only in Eukarya and likely evolved after the divergence with Archaea. (**Right**) Distribution of RNase Ps in Eukarya using the five supergroup model (aspects of the branching order remain controversial) (12, 13). PRORPs are found in four supergroups: Excavata, Archaeplastida, SAR [Stramenopiles, Aveolata, and Rhizaria], and Opisthokonta, but not in the fifth, Amoebozoa (11). Dashed blue or yellow lines indicate that some clades within the supergroup lack RNase P RNA or PRORP sequences, respectively.

RNA-Based RNase Ps Exist in Bacteria, Archaea, and Many Eukarya

RNA-based RNase Ps are ribonucleoprotein (RNP) complexes containing one conserved catalytic RNA subunit and a variable number of protein subunits depending on the organism. In general, protein content increases from ~ 10% to ~

40% to ≥ 70% for the enzymes from Bacteria to Archaea to Eukarya, respectively. Unlike self-cleaving ribozymes, RNA-based RNase Ps catalyze multiple turnovers and remain unchanged after catalysis. Bacterial, archaeal, and eukaryotic nuclear RNase P RNAs share a core consensus sequence with a conserved secondary structure (14). There are three distinct types of secondary structures in Bacteria: type A (ancestral), B (*Bacillus*), and C (Chloroflexi), and two in Archaea: M (Methanococci) and P (Pyrobaculum), while the eukaryotic nuclear RNase P RNA has a larger variety of secondary structures (15-20). The secondary structure of a type A RNase P RNA from *Thermotoga maritima* is shown in Fig. 1-3.

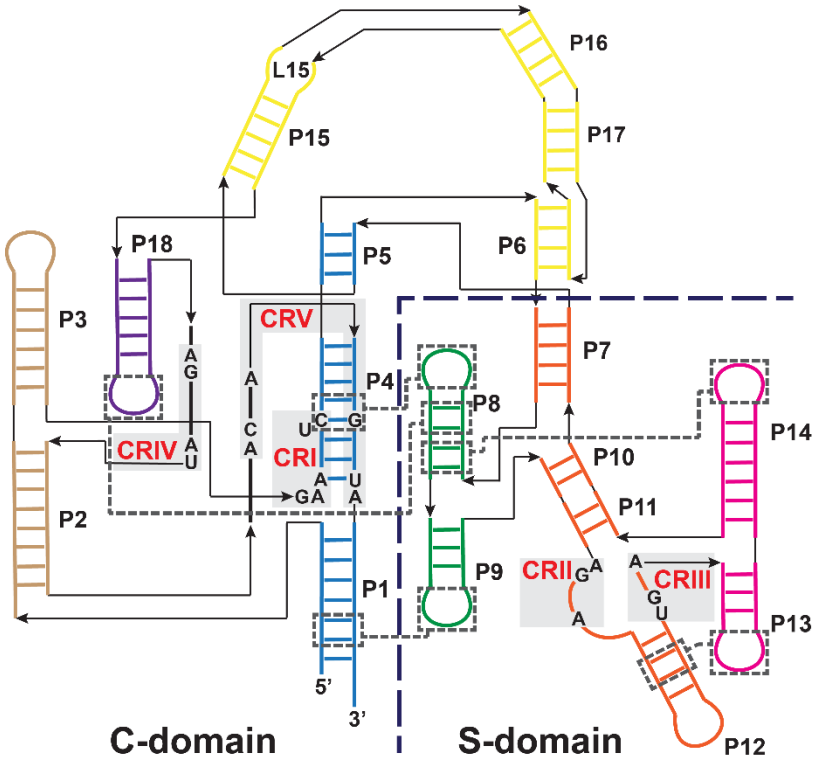


Figure 1-3: The secondary structure of type A RNase P RNA from *T. maritima* (adapted with permission from (21)) (22, 23). Individually-folding catalytic (C)- and specificity (S)-domains are divided by the dashed line. Conserved regions (CRI–V, red) are indicated by gray shading and numbered in order of occurrence from the 5' end (4). Tertiary interactions in the *T. maritima* P RNA are indicated by dashed gray boxes and lines. Helices are colored by coaxial-stack and are numbered as P1–P18 in order of occurrence from the 5' end. The C-domain includes P1/P4/P5 (blue), P2/P3 (brown), P6/P15/P16/P17 (yellow) and P18 (purple). The S-domain includes P7/P10/P11/P12 (orange), P8/P9 (green), and P13/P14 (pink).

RNase P in Bacteria is composed of one large RNA subunit (P RNA), typically 340–420 nucleotides, and one small protein subunit (RnpA), approximately 14 kDa (24). Both subunits of the bacterial ribozyme are essential for efficient catalytic activity *in vivo*, while the RNA subunit is sufficient for catalysis *in vitro* at high cation

concentrations (3). The RnpA protein is approximately 10% of the bacterial holoenzyme by mass and enhances the affinity of RNase P for pre-tRNA and metal ions (25-28). RnpA contacts the leader sequence of pre-tRNA to stabilize an active enzyme-substrate conformer (29).

Humans and other metazoans have two RNase P enzymes: a nuclear RNA-based RNase P and a mitochondrial PRORP. Human nuclear RNase P contains a single RNA subunit and at least ten protein subunits, seven of which are homologous to the protein subunits in yeast (Table 1) (5, 30, 31). The human mitochondrial PRORP consists of three nuclear-encoded RNase P proteins (MRPP1, 2, and 3), none of which are homologous to the RNA-based RNase P proteins (see following section) (32).

Table 1-1: Conservation of RNase P protein components across the domains of life. RnpA proteins from various Bacteria share homology to each other, but not to archaeal or eukaryotic proteins. Archaeal and eukaryotic nuclear RNase P proteins share sequence homology with proteins on the same row. Protein-only RNase P components do not share homology with any of the RNA-based RNase P protein components.

		Eukarya			
Bacteria	**Archaea** [a]	**RNA-based**		**PRORP**	
		Yeast	**Human**	**Human**	**Plant**
RnpA					
	Mth687p	POP5	hPOP5		
	Mth11p	POP4	RPP29		
	Mth688p	RPP1	RPP30		
	Mth1618p	RPR2	RPP21		
	L7Ae [b]	POP3	RPP38		
		POP1	hPOP1		
		POP6	RPP25		
		POP7	RPP20		
		POP8			
			RPP40		
				MRPP1	
				MRPP2	
				MRPP3	PRORP1–3

[a] Naming for RNase P proteins from *Methanobacter thermoautotrophicus* (31).
[b] L7Ae homologs are only associated with M-type archaeal RNase P (33).

PRORPs are Found Only in Eukarya

PRORP sequences are found broadly in the nuclear genomes of Eukarya, including animals (Metazoa), plants and algae (Archaeplastida), trypanosomes

(Excavata), and heterokonts (SAR subgroup Stremenopiles) (Fig. 1-2) (see (11) for a more complete discussion). PRORPs from the land plant *Arabidopsis thaliana* are able to complement viability in an *E*. *coli* strain with the RNase P RNA under control of the arabinose promoter when grown on glucose media, as well as a yeast strain lacking core protein components for the RNA-based RNase P, indicating that PRORPs can maintain the essential biological functions of the RNA-based RNase Ps (34, 35).

Human mitochondrial RNase P (mtRNase P) was the first RNase P enzyme systematically shown to lack an RNA component (32). This enzyme requires three MRPP subunits: tRNA methyltransferase 10 C (TRMT10C; MRPP1), a short-chain dehydrogenase/reductase (SDR5C1; MRPP2), and a novel metallonuclease, which houses the RNase P active site (huPRORP; MRPP3). All three subunits were affinity-purified and all were required to reconstitute efficient pre-tRNA cleavage activity *in vitro* (32). SiRNA knockdowns of either MRPP1 or MRPP3 resulted in precursor transcript accumulation in mitochondria from HeLa cells (36).

MRPP3 represents a novel class of endonucleases, containing an N-terminal pentatricopeptide repeat (PPR) RNA-binding domain, a central structural Zn^{2+}-binding domain, and a Nedd4-BP1, YacP nuclease (NYN) domain with homology to the flap endonucleases (37). The molecular basis for the requirement of MRPP1 and MRPP2 for metazoan mtRNase P activity remains to be clarified. MRPP1 is one of three human homologs of Trm10, a SAM-dependent methyltransferase that catalyzes the addition of a methyl group to the *N*1 of G9 in yeast tRNAs (38). MRPP1 and MRPP2 form a stable methyltransferase complex (MRPP1/2) which catalyzes the addition of a methyl group to the *N*1 of a purine (A or G) at position 9 in mitochondrial tRNAs (39). Both pre-tRNA and mature tRNA are substrates for the MRPP1/2 complex (39). However, this methyltransferase activity is not required for RNase P catalysis. The MRPP1/MRPP2 complex is proposed to enhance the affinity of MRPP3 for the substrate by protein–protein contacts, by reorganizing the tRNA structure for recognition by MRPP3 and/or for altering the structure of the MRPP3 active site (32, 39, 40).

Plants, algae, and trypanosomes encode homologs of MRPP3 that do not require additional subunits for catalysis *in vitro*. The algae, *O*. *tauri*, encodes one PRORP homolog that catalyzes pre-tRNA processing *in vitro* (41), trypanosomes encode two functional PRORP paralogs (42), and land plants encode three functional PRORP paralogs (34).

In *A*. *thaliana* (*At*), PRORP1 is localized to the mitochondria and chloroplasts, while PRORP2 and PRORP3 co-localize to the nucleus (34). *At*PRORP2 and *At*PRORP3 have overlapping substrate selectivity, since depletion of both enzymes is required to eliminate RNase P activity in the nucleus (43). They are highly similar, sharing 80% sequence identity and 88% similarity, suggesting a relatively recent gene duplication event. Given that it is not possible to obtain viable cells containing a deletion of PRORP1 or a double deletion of PRORP2/3, it was proposed that *A*. *thaliana* is devoid of an RNA-based RNase P (43).

STRUCTURES OF RNase Ps

There have been a variety of structures solved for both RNA-based RNase Ps and PRORPs from all three domains of life including individual subunits, subcomplexes, and holocomplexes. In addition to visualizing features previously proposed from biochemical experiments, many structures have provided significant insight into structural features that are key for catalysis and substrate recognition.

Bacterial RNase P Ribozyme

The highly conserved core structure of bacterial RNase P RNAs includes two independently folding domains: the C- and S-domains (Fig. 1-3). The C-domain comprises 60% of the P RNA and can catalyze pre-tRNA cleavage both in the presence and absence of the RnpA protein (44-46). The S-domain comprises the remaining 40% of the P RNA and is important for substrate affinity/specificity, interacting with the D- and TψC-domains of pre-tRNA (47-49). A crystal structure of the *T*. *maritima* RNase P holoenzyme, in a product complex with a yeast $tRNA^{Phe}$ and a 5’ leader, revealed important structural information about the ternary complex (PDB 3q1r, Fig. 1-4A) (22). The P RNA S-domain interacts with the D-

and TψC-loops of tRNA, while non-helical elements in the C-domain contribute to interactions with the minor groove of the tRNA acceptor-stem.

Two metal-binding sites located in the active site near the P4 helix, previously identified through biochemical techniques, were validated by soaking the crystals with Eu^{2+} and Sm^{3+} (Fig. 1-4B) (22). Metal 1 is located between the tRNA 5' end, the phosphodiester bonds of A_{50} and G_{51} and the carbonyl oxygen O4 of U_{52}, which is universally conserved. Metal 2 is located close to the phosphoryl oxygens of G_{51}, the 3' OH of the leader, and the 5' end of tRNA.

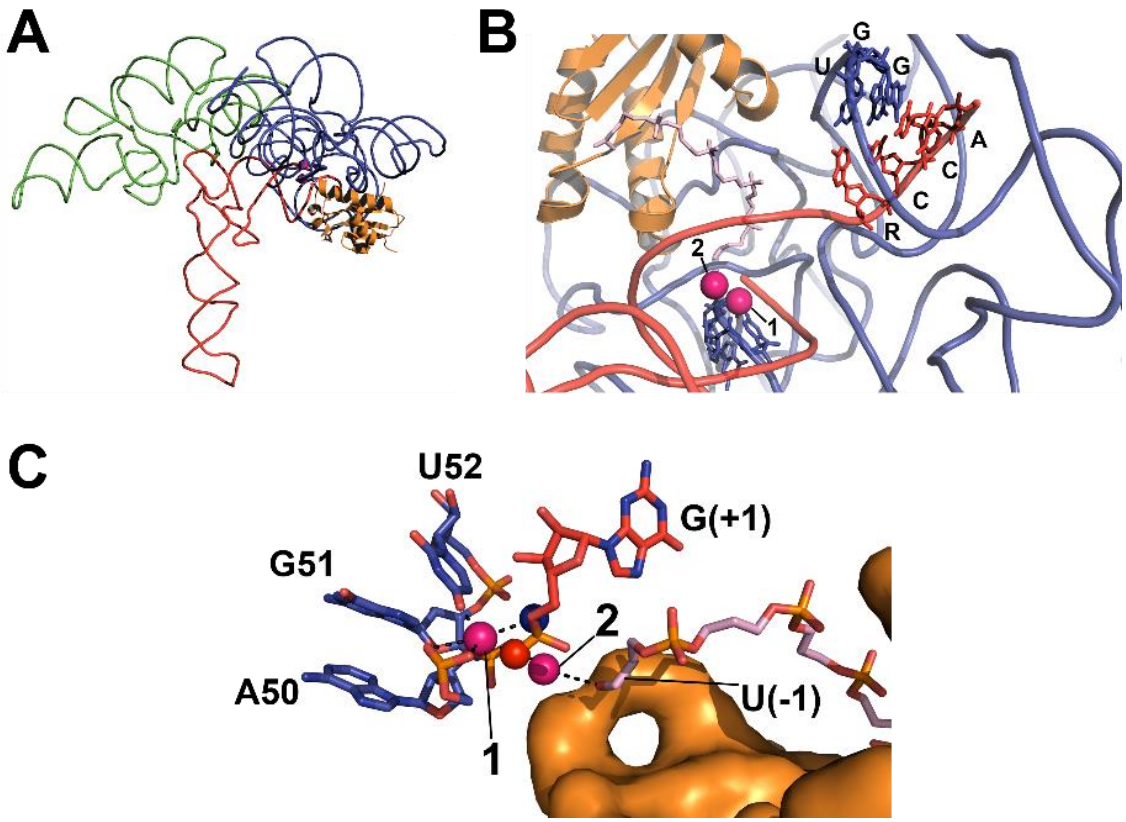


Figure 1-4: X-ray crystal structure of *T. maritima* RNase P-tRNA-leader product complex (PDB 3q1r, PyMOL) (22). (**A**) The holoenzyme-product complex of RNase P is shown, including C- (blue backbone) and S- (green backbone) domains, tRNA (red backbone), and RnpA (orange cartoon); (**B**) Topology of substrate contact sites in the catalytic domain (colored as in A), including the 5' leader (light pink sticks) bound to RnpA and proposed divalent metal ions (pink spheres). Base-pairing between GGU residues in P RNA and the 3' RCCA of tRNA (U_{256}–R_{73}, G_{255}–C_{74}, G_{254}–C_{75}) is shown; (**C**) Topology of the active site (colored as in A), including the active site residues (blue carbon atoms), product G_1 (red carbon atoms), the 5' leader (light pink carbon atoms) bound to RnpA (orange surface), proposed metal contacts (black dashed lines), and positions of the pro-R_P (red sphere) and pro-S_P (blue sphere) oxygens of the product 5' phosphate.

Protein-Only RNase Ps

X-ray crystal structures have been solved for both the single-subunit plant PRORPs and the multi-subunit metazoan PRORPs. In the single-subunit plant enzymes, the structures of *At*PRORP1 and *At*PRORP2 visualized similar folds (37, 50). Two structures of huPRORP/MRPP3 with extensive truncations in the PPR domains were reported (40, 51). Structures of MRPP2, which is a homotetramer in solution and *in crystallo*, have been solved (52). No structures of MRPP1 have been solved, but there are several partial structures of homologous Trm10 family tRNA methyltransferases visualizing the methyltransferase domain (53).

The crystal structure of *A. thaliana* PRORP1 revealed a three-domain architecture, with an NYN metallonuclease domain, a bipartite central Zn^{2+}-binding domain, and a PPR RNA-binding domain (Fig. 1-5A) (37). The active site, located within the NYN domain, contains four aspartates (*At*PRORP1 Asp 399, 474, 475, and 493) that are fully conserved across the available PRORP sequences, with a fifth aspartate (*At*PRORP1 Asp 497) not conserved in some metazoan homologs. The structure of *At*PRORP1 bound to Mn^{2+} (PDB 4g24), which activates the enzyme, revealed density for two metal ions bound in the active site (37). Mechanistic studies with *At*PRORP1 provided further evidence for a two-metal ion mechanism (54), which is discussed in the following section.

The crystal structure of *At*PRORP2 revealed a more “open” conformation, in which the active site and PPR domain are rotated away from one another (Fig. 1-5A) (50). Interestingly, the PRORP2 monomers pack in an extended chain, with a conserved lysine (*At*PRORP2 Lys 42) from one chain inserted into the active site of the neighboring chain (PDB 5diz). While no metal densities were obtained by soaking metals into the crystal, the construct was active and monomeric in solution (50). Thus, the dimerization interactions *in crystallo* are not likely to be relevant to *in vivo* function.

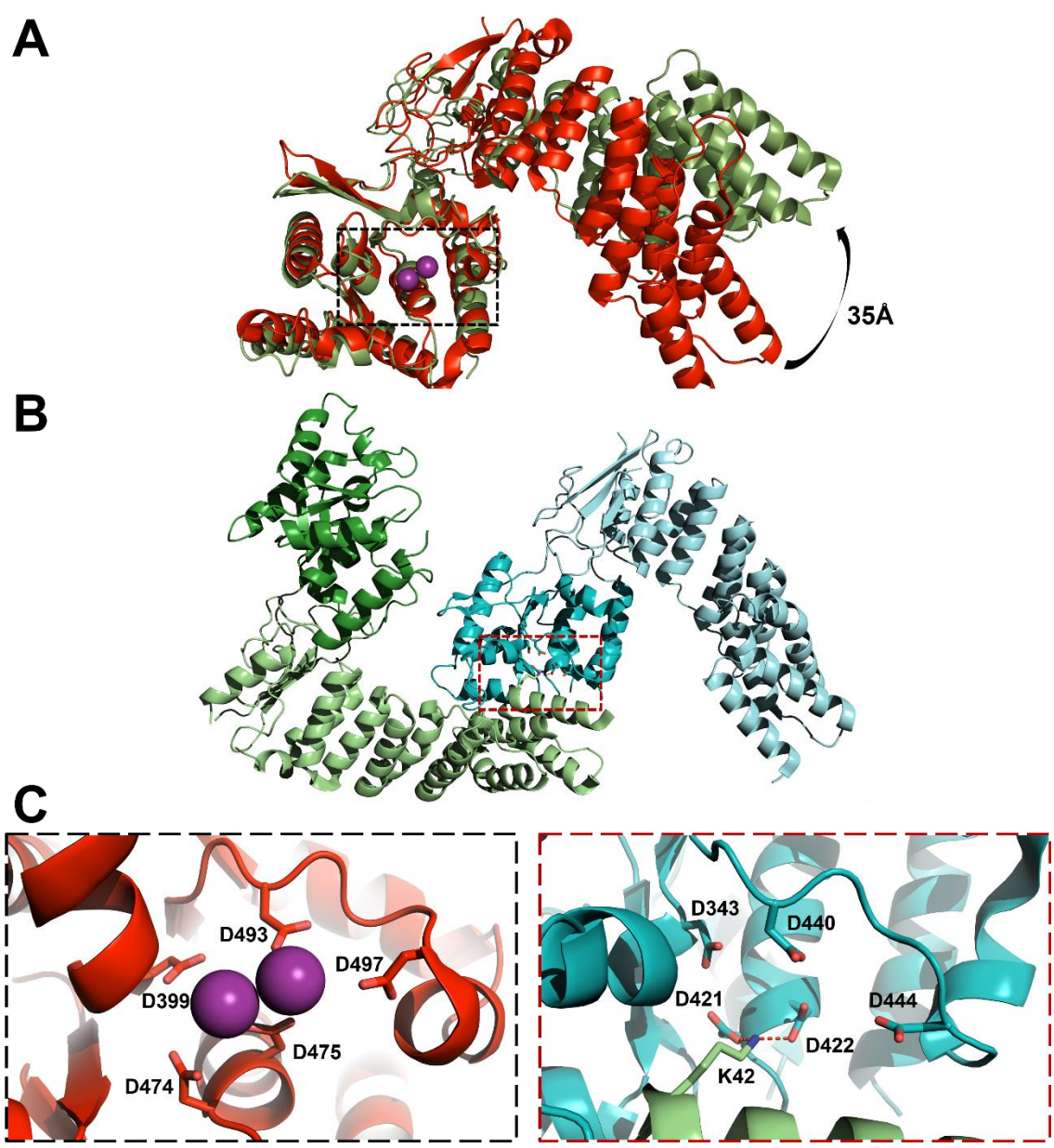


Figure 1-5: Crystal structures of *At*PRORP1 (PDB 4g24) and *At*PRORP2 (PDB 5diz) (37, 50). (**A**) Three-dimensional alignment of *At*PRORP1 and *At*PRORP2 structures via their active sites (dashed box). PRORP1 (red) topology includes the NYN metallonuclease domain bound to Mn^{2+} (purple spheres). PRORP2 (green) is in a more "open" conformation, resulting in a ~ 35 Å difference in the position of the first PPR helix; (**B**) *At*PRORP2 crystallization dimer (left), with subunits labeled in green and cyan (NYN domains in brighter colors); (**C**) Expanded views of the dashed boxes in panel A (left) or B (right). (**Left**) The AtPRORP1 active site includes four fully-conserved aspartates and one partially-conserved aspartate (sticks). (**Right**) Close-up of intersubunit interaction between *At*PRORP2 molecules, with active site aspartates of one molecule interacting with Lys 42 (sticks) of the second (red dashed lines).

Two recent crystal structures revealed a V-shaped structure for MRPP3 comparable to *At*PRORP1 (PDB 4xgl, 4rou; Fig. 1-6) (40, 51). The MRPP3 constructs used in both structures include truncation of two or four of the PPR

motifs. Both structures contain significant disorder in the NYN domain, the metal 1 site is occluded by asparagine 412, and arginine 445 distorts the positions of Asp 478–Asp 479 (Fig. 1-6, middle/right) (40, 51). The structure reported by Reinhard, et al. (40) also visualizes Arg 498, which forms a hydrogen bond with Asp 499–the equivalent of the metal ligand Asp 399 in *At*PRORP1–and occludes the metal 2 site (Fig. 1-6, middle) (40).

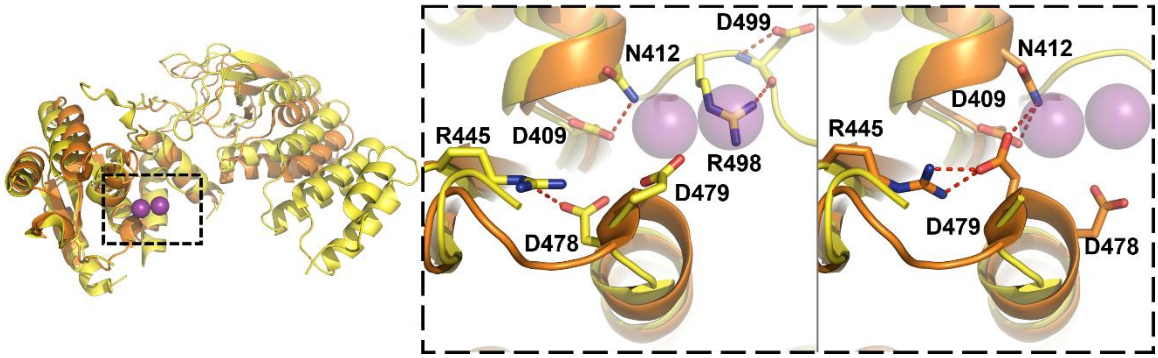


Figure 1-6: Alignment of two MRPP3 structures (PDB 4xgl, yellow; PDB 4rou, orange) with the *At*PRORP1 active site (PDB 4g24, only metal ions shown) (37, 40, 51). (**Left**) Alignment of two MRPP3 structures with the *At*PRORP1 active site (purple spheres). MRPP3 NYN residues (4xgl: Ser 361/Pro 362, Asp 409, Asp 479, and Asp 499, or 4rou: residues Ser 361/Pro 362, Asp 409, Asp 478, and Asp 479) were aligned to the equivalent *At*PRORP1 residues. The dashed box is expanded on the right. (**Middle** and **Right)** Dashed box contains the aligned MRPP3 active site residues from 4xgl (middle, yellow) or 4rou (right, orange) with Mn^{2+} from *At*PRORP1 (transparent purple spheres). Residues with the potential to occlude the metal sites, Asn 412, Arg 445, and Arg 498, are shown where coordinates are available. Hydrogen bonds between residues are shown as dashed red lines.

The authors independently ascribe the NYN domain disorder and active site occlusion as an inactive conformation that prevents metal binding, and suggest that the MRPP1/2 complex activates catalysis by promoting an active conformation (40, 51). However, even the more complete MRPP3 construct was inactive in assays containing MRPP1/2 (40), so the distorted active sites might also be related to the N-terminal deletion. Similar deletions of PPR motifs in PRORP1 also inactivated the enzyme (37, 55). Furthermore, the active site of *At*PRORP2 also does not contain metal density and makes interactions with a neighboring molecule (Fig. 1-5B) (50), similar to the active site of MRPP3 (51). In the case of *At*PRORP2 this is considered an artifact of crystallization, as *At*PRORP2 has activity comparable to *At*PRORP1 *in vitro* (50).

Both groups attempted domain swaps with *At*PRORP1, yet exchanging the NYN domain did not restore activity for MRPP3 alone and was detrimental to the

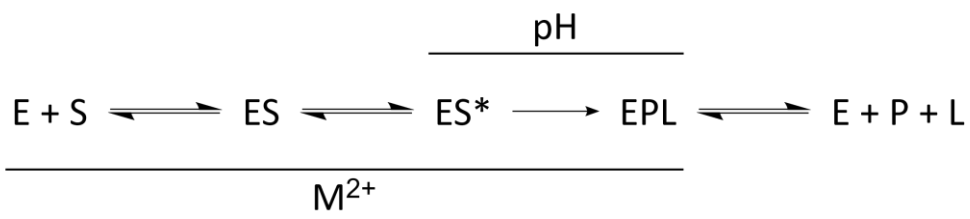
activation by MRPP1/2 (40). A larger swap, including the central domain and the final helices of the PPR domain, restored some cleavage activity in the absence of MRPP1/2 (51). The structural and NYN sequence similarities between PRORP1 and MRPP3 probably indicate similar catalytic mechanisms, though more information on MRPP3 structure and catalysis is required to allow a full comparison.

CATALYSIS BY RNase Ps

RNase P enzymes are metal ion-dependent endonucleases. More specifically, they are hydrolases that catalyze site-specific phosphodiester bond hydrolysis primarily within pre-tRNA. Research into the mechanisms of RNase P catalysis, including the modes of nucleophile activation, have provided insight into the diverse evolutionary backgrounds capable of performing this fundamental biological reaction.

Kinetic Mechanism

Transient kinetic studies of *B*. *subtilis* RNase P established a minimal four-step kinetic mechanism (Scheme 1) (28, 56). In this mechanism, RNase P (E) and pre-tRNA (S) associate in a two-step binding event in which they first form an enzyme–substrate complex (ES) in a near diffusion-limited binding step. Once bound, the ES complex isomerizes to a catalytically-competent conformer (ES*) in a metal-dependent step, referred to as the “conformational change” step. Pre-tRNA is then cleaved to form mature tRNA and 5’ leader and the products dissociate. Product release is rate-limiting for the bacterial enzyme, resulting in a kinetic burst under multiple turnover conditions (57).



Scheme 1-1: Minimal kinetic mechanism of bacterial RNase P ribozyme catalysis. The initial binding step is bimolecular and dependent on the concentrations of both E and S. The rates of binding, conformational change, and substrate hydrolysis are all dependent on divalent metal ions (M^{2+}), while hydrolysis is also dependent on pH.

At low pH and under single turnover conditions, the observed rate constant (k_{obs}) increases with a log-linear relationship to pH (56). The Hill coefficient, n_H = 1, is consistent with a single ionization producing the metal-hydroxide nucleophile required for catalysis (56). This is in contrast to nucleases that utilize a general base, which display n_H = 2 (58, 59). The conformational change step becomes rate-limiting at high pH, resulting in a kinetic (rather than thermodynamic) pK_a (56). In the two-step binding mechanism, ES* is stabilized by at least two classes of inner-sphere metal ions (28). A class of high-affinity divalent cations is required to stabilize the ES* conformer, while a lower-affinity metal ion activates catalysis (26, 28, 29, 60). The conformational change could allow the C- and S-domains to position functional groups and catalytically important ions in the active site (61), as well as position the active site with respect to substrate for catalysis (62, 63). Of note, the crystal structure of the *T*. *maritima* RNase P may not reflect the active ES* conformation as it is a product complex and does not contain pre-tRNA (22).

RNase P Ribozyme Metal-Binding Sites

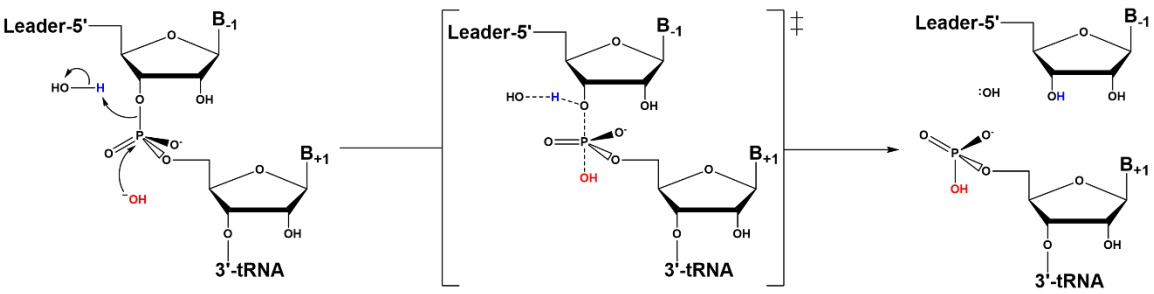
Identifying the position of metal ions involved in catalysis in RNase P RNA is challenging because the majority of divalent metal ions bind nonspecifically via electrostatic interactions (64). Additionally, P RNA requires divalent ions for stabilizing the folded structure (65-67). Further, crosslinking studies of *E*. *coli* RNase P RNA examining the position of the pre-tRNA cleavage site relative to the P4 helix suggest that metal ion-binding in the P4 helix indirectly stabilizes catalytic metal ions that interact with the scissile phosphodiester bond (68).

Important early work from Pace and colleagues identified specific sites on the P4 helix of the bacterial P RNA that are important for catalysis (69, 70). Specifically, rescue of phosphorothioate substitutions with Mn^{2+} or Cd^{2+} revealed that the non-bridging phosphodiester oxygens of P4 helix residues A_{50} and G_{51} coordinate to the catalytic metal ions through inner-sphere interaction (71). These experiments suggested metal coordination through the pro-S_P oxygen of A_{50} (in the P RNA but not the holoenzyme), and to both pro-R_P and pro-S_P oxygens of G_{51} (71, 72).

Additional catalytic metal sites were identified on the pre-tRNA substrate. Both *R*- and *S*-phosphorothioate substitutions at the scissile phosphodiester bond of the pre-tRNA substrate decrease the cleavage rate constant catalyzed by RNase P RNA by 10^3–10^4-fold. The activity of the *R*-phosphorothioate-substituted pre-tRNA can be partially rescued by addition of thiophilic metal ions (Mn^{2+} and Cd^{2+}) (73-77). Such rescue experiments provided evidence for inner-sphere coordination of two metal ions to the pro-R_P oxygen of pre-tRNA. At the N_{-1} position of pre-tRNA, 2'-NH_2 and 2'-H substitutions decrease both the catalytic rate constant and the Mg^{2+} affinity for RNase P RNA. Mn^{2+} can rescue cleavage of both 2'-NH_2 and 2'-H pre-tRNA, suggesting that the 2'-OH of pre-tRNA at the cleavage site also interacts with a metal-bound water molecule (77).

Hydrolysis Requires Activation of a Metal-Bound Water Molecule

Nucleophile activation in RNase P catalysis is proposed to proceed by stabilizing a metal-bound hydroxide. Kinetic isotope effect studies on RNase P RNA catalysis conducted by the Harris group suggested that a metal-bound hydroxide, which is not deprotonated by a general base, serves as the nucleophile in the reaction (Scheme 2) (78). The isotope effect of the reaction of RNase P with pre-tRNA in 50% ^{18}O-labeled water was compared to that of two model systems: 1) hydrolysis of thymidine 5' *p*-nitrophenyl monophosphate (T5PNP) catalyzed by Mg^{2+}, which occurs via a concerted mechanism with a pentacoordinate transition state; and 2) conversion of adenosine to inosine catalyzed by adenosine deaminase (ADA), which uses a stepwise mechanism with formation of a tetrahedral intermediate (78). The data indicated that the bonding environment in the transition state of RNase P was more akin to that of T5PNP hydrolysis (78, 79).



Scheme 1-2: The concerted hydrolytic mechanism proposed for bacterial RNase P. The acid may be a water or metal-bound water.

RNase Ps Utilize Similar Catalytic Strategies

Magnesium ions (Mg^{2+}) activate RNase P catalysis *in vitro* and *in vivo* (3, 67). The *T*. *maritima* RNase P product complex crystal structure visualized two potential metal sites (22), yet the number of metals involved in the active site chemistry remains in question. Mechanistic models with three activating metal ions have also been proposed, including a co-catalytic metal bound near the 2'-OH on N_{-1} of the substrate The general two-metal ion mechanism proposed for nucleases and polymerases over 20 years ago has largely been vindicated by subsequent studies and is consistent with the PRORP mechanism (54, 80-82). The general two-metal ion mechanism includes metal 1 positioning and activating a hydroxide nucleophile, while metal 2 stabilizes the transition state and coordinates a water molecule to protonate the 3' oxyanion leaving group (22, 74, 75, 80, 83). Additionally, both metal ions are proposed to stabilize the developing charge in the transition state.

PRORP1 from *A*. *thaliana* is the PRORP best described mechanistically. It has been used as a model system to study mammalian mitochondrial RNase P, given that it is homologous to the nuclease subunit of the human enzyme. The mechanistic data available enables a detailed comparison between the protein-only and the bacterial ribozyme RNase P enzymes. In contrast to the ribozyme, data suggest that the active site metals of PRORP1 do not contact the pro-R_P oxygen of the scissile phosphate (84). Experiments with substrates containing an *R*-phosphorothioate at the scissile phosphate indicated only a five-fold reduction in Mg^{2+}-dependent PRORP activity, compared to the > 26000-fold decrease for the ribozyme (84). In PRORP the pro-S_P oxygen of the scissile bond of pre-tRNA is proposed to contact a metal ion based on the absence of a pro-R_P effect and by homology to other nucleases (54, 84, 85).

The difference in metal coordination at the scissile phosphodiester may be explained by altered stereochemical requirements for coordinating the pro-R_P or pro-S_P oxygen atoms (Fig. 1-7). In the canonical cloverleaf structure, the pro-R_P oxygen of the pre-tRNA substrate faces inwards towards the minor groove.

Therefore, metal coordination of this oxygen in the ribozyme active site requires insertion of P RNA into the minor groove of the tRNA acceptor stem (Fig. 1-7, top). The result is that the groove is ≥ 5 Å wider in the ribozyme-bound structure than in a typical tRNA (Fig. 1-7, compare top right to bottom right). The PRORP active site is relatively flat, such that a similar insertion to position the metals near the pro-R_P oxygen would require large changes in the NYN structure and/or extensive distortions to the pre-tRNA acceptor helix. The small effects of the *R*-phosphorothioate substitution at the scissile bond of pre-tRNA argue against these distortions in a PRORP-substrate complex (84).

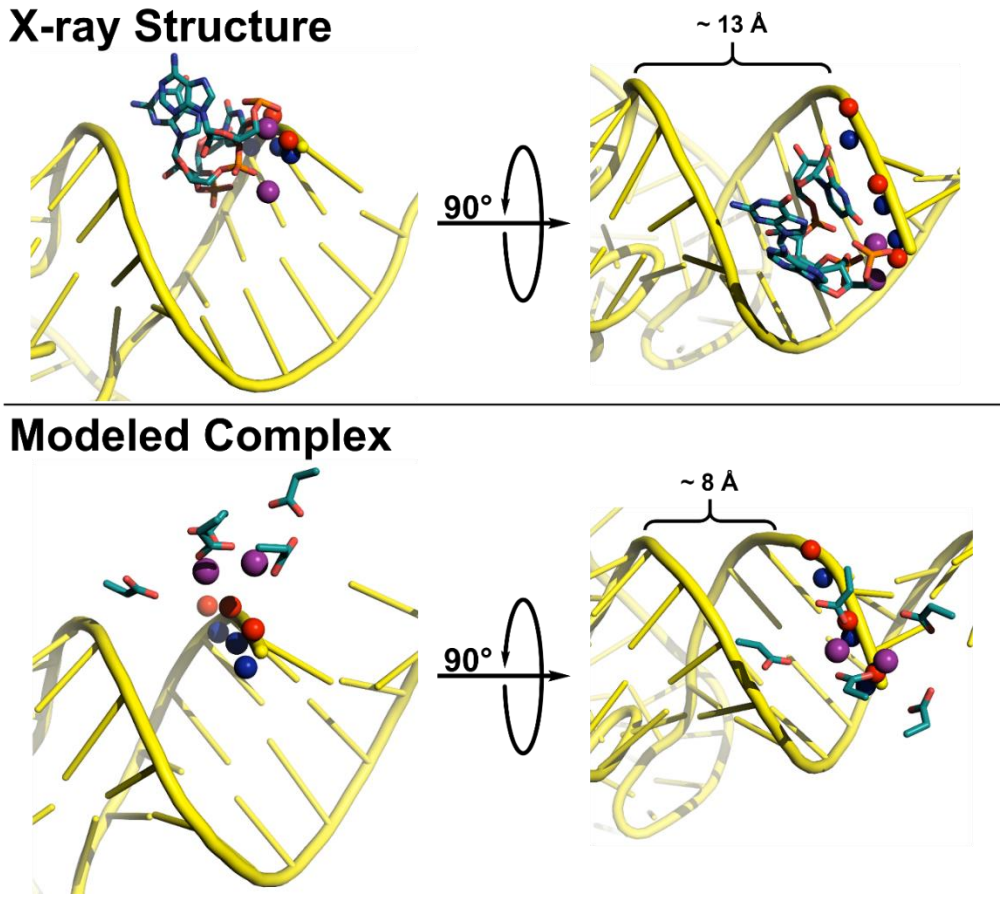


Figure 1-7: Active site coordination of substrate by *T. maritima* RNase P (top, x-ray crystal structure) and *At*PRORP1 (bottom, complex modeled in PyMOL). The minor groove width was measured as the distance between the non-bridging phosphate oxygens. (**Top**) *T. maritima* product complex crystal structure (PDB 3q1r) (22). Pro-R_P (blue spheres) and pro-S_P (red spheres) oxygen atoms of tRNA product (yellow cartoon) shown for N_1–N_3. Active site metal atoms (purple spheres) and metal-coordinating residues A_{50}, G_{51}, and U_{52} are visualized (teal sticks). (**Bottom**) The PRORP1 active site (PDB 4g24) was aligned to *S. cerevisiae* tRNAAsp (PDB 2tra) using the human DNA exonuclease I active site bound to DNA (PDB 3qeb) as a guide (37, 86, 87). The tRNA, backbone oxygen atoms, active site metal atoms, and active site residues Asp 399, Asp 474/475, Asp 493, and Asp 497 are colored as in the top panel.

Like the ribozyme, a single ionization (n_H = 1) with increased activity at high pH is observed in the pH-dependence of *At*PRORP1 (54, 56). Furthermore, no decrease at high pH was observed in the range tested, indicating that no group with a pK_a ≤ 9.5 acts as a general acid (54). This might be surprising, given that a frequent component of the RNA world hypothesis posits that protein replaced RNA as the catalytic molecule due to the larger suite of side chains capable of participating in chemistry under biological conditions (10). Alanine mutations of four conserved aspartates (Asp 399, Asp 474, Asp 475, and Asp 493) significantly reduced activity (37). Two of the mutants (D474A and D475A) were rescued by increasing the metal concentration (54). The other two mutants (D399A and D493A) were not rescued, which may indicate they have catalytically-important functions in addition to metal binding (54), though these have not been elucidated. Mutation of Asp 497 has not been attempted, but in the crystal this side chain is positioned to form an outer sphere interaction with metal 2 (54).

The catalytic efficiencies achieved by PRORP enzymes (43, 54, 88, 89) are comparable to or as much as 10^3-fold lower than those reported for RNA-based RNase Ps (25, 90), inconsistent with an RNA world model in which protein enzymes evolved due only to catalytic enhancements over their RNA world predecessors. In the case of RNase P, it appears that both the RNA-based and protein-only enzymes function to catalyze phosphodiester bond hydrolysis mainly by correctly positioning hydrated metal ions at the cleavage site (Fig. 1-8).

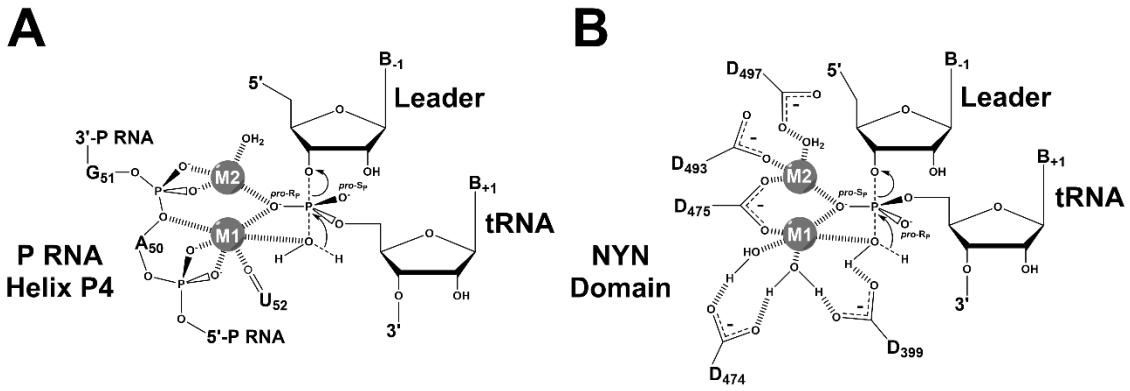


Figure 1-8: Proposed transition state structure of the catalytic mechanisms of (**A**) bacterial RNase P (as proposed in (22)) and (**B**) *At*PRORP1 (adapted from (54)).

SUBSTRATE RECOGNITION BY RNase Ps

Most tRNAs from all domains of life have a cloverleaf secondary structure and an L-shaped tertiary structure (Fig. 1-9) (91). In metazoan mitochondria, many tRNAs deviate from the canonical primary and secondary structure and these tRNAs can be divided into four groups (92). One type is similar to canonical tRNAs with conserved tertiary interactions, while the other three groups have truncated structural elements or entirely lack certain structural features, resulting in the loss of conserved interactions. These non-canonical tRNA structures led to the proposal that the additional subunits of human mtRNase P evolved to recognize these substrates (93).

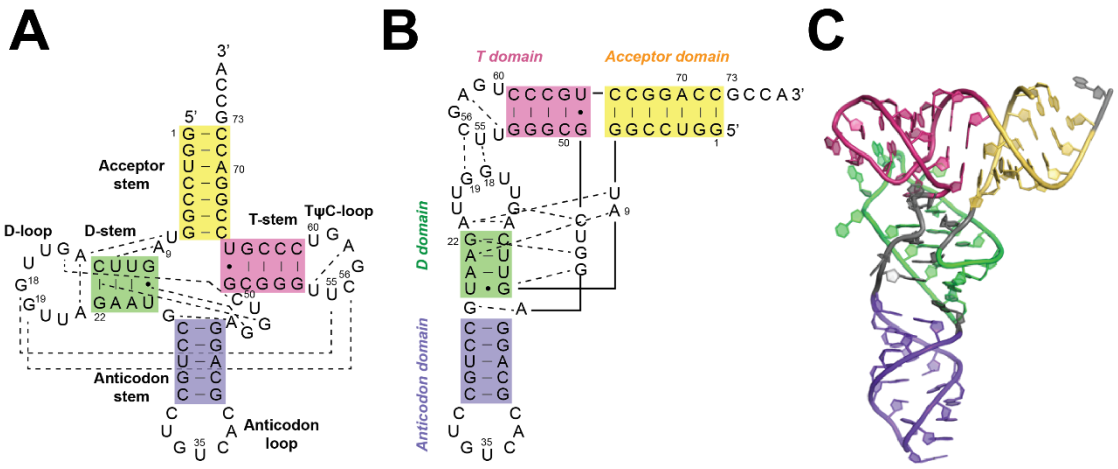


Figure 1-9: Secondary and tertiary structures of canonical tRNA (adapted with permission from (7)). (**A**) and (**B**) Secondary structure of *B*. *subtilis* $tRNA^{Asp}$ with tertiary interactions denoted by dashed lines. (**C**) A crystal structure of an unmodified *E*. *coli* $tRNA^{Phe}$ (PDB 3l0u) (94).

Recognition by Bacterial RNase P

Both the RnpA protein and P RNA subunits contribute to recognition of pre-tRNA substrates. The holoenzyme has a higher binding affinity for pre-tRNA relative to tRNA (25). The holoenzyme affinity for pre-tRNA varies with the leader length from 2- to 5-nt (60, 95). This trend was not observed for P RNA alone (60, 95), which provided evidence that RnpA enhances substrate affinity and cleavage activity of RNase P by interacting with the 5' leader. Interactions with the 5' leader also increase the affinities of Mg^{2+} ions bound to the RNase P–pre-tRNA complex and result in uniform binding affinity for various pre-tRNA substrates by combining

a low affinity tRNA body with a high affinity leader sequence (26, 96). Cross-linking and fluorescence resonance energy transfer (FRET) experiments first demonstrated a direct interaction between the 5' leader and RnpA (95, 97), and the leader position observed in the *T*. *maritima* RNase P complex structure is consistent with these biochemical data (22).

Sequence-specific interactions between bacterial RNase P and substrates have been identified in both the 5' leader sequence flanking the tRNA genes and the 3' RCCA motif (Fig. 1-10) (98-104). These interactions enhance pre-tRNA affinity and cleavage. The 3' RCCA motif of tRNA forms Watson-Crick base pairs with a GGU sequence in P RNA, which enhances pre-tRNA and Mg^{2+} affinity of P RNA and to a lesser extent of the holoenzyme (105-109).

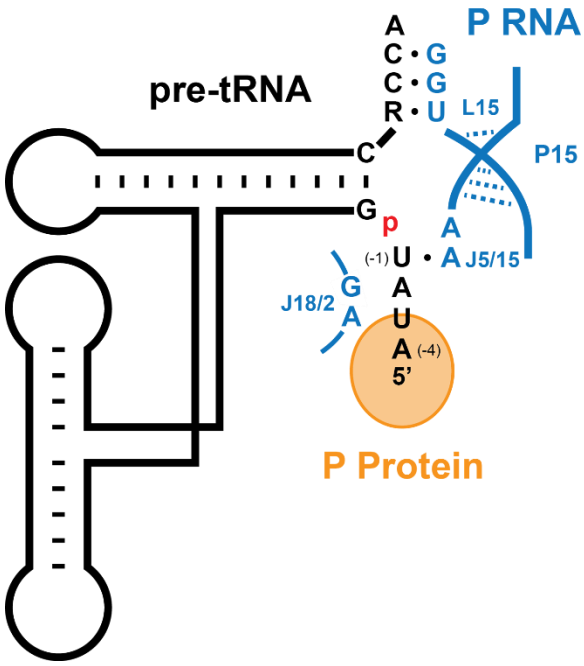


Figure 10: Sequence specific interactions between 5' and 3' sequences of pre-tRNA and bacterial RNase P (adapted with permission from (7)). The 3' RCCA and U(-1) base pair with the GGU motif in the L15 loop and A213, respectively, while A(-2) makes a non-Watson-Crick interaction with U294 (23, 101, 110).

In both *E*. *coli* and *B*. *subtilis* RNase Ps, biochemical studies identified a sequence preference for uracil at position N_{-1} in the 5' leader of pre-tRNA. Base pairing between this nucleotide and an adenosine in P RNA is proposed (101, 110). Statistical analyses of the sequences of pre-tRNA genes for these species also reveal a preference for U_{-1} (104). *B*. *subtilis* RNase P displays a preference

for A_{-2}, forming a trans Watson-Crick–sugar edge interaction with a P RNA guanosine (23).

The D–TψC interaction is an important factor for pre-tRNA recognition and cleavage site selection by bacterial RNase P. The *E*. *coli* RNase P holoenzyme can tolerate model hairpin-loop substrates that mimic the acceptor and TψC-stems, as well as shorter three base pair hairpin loops, but these are miscleaved at a higher frequency than more complete substrates (63, 111). Deletion or elongation of the D-stem of a cyanobacterial pre-$tRNA^{Gln}$ substrate led to considerable miscleavage at N_{-1} by the *E*. *coli* RNase P holoenzyme (112). These results are consistent with structural data indicating interactions between the S-domain and the D–TψC loops (22).

Recognition by Protein-Only RNase Ps

Most plant tRNAs have canonical secondary and tertiary structures (113, 114), so it is not unexpected that *At*PRORPs efficiently cleave bacterial substrates such as *Thermus thermophilus* pre-$tRNA^{Gly}$ (84) and *B*. *subtilis* pre-$tRNA^{Asp}$ (54). *At*PRORP2 also cleaves the *T*. *thermophiles* pre-$tRNA^{Gly}$ bacterial substrate at 28°C, but not at 37°C (84), which suggests that structural dynamics of either pre-tRNA or *At*PRORP2 are important for recognition and catalytic efficiency. Furthermore, substrate recognition strategies of the three *At*PRORP isozymes are similar, as catalytic efficiencies with four pre-tRNA substrates varied < 10-fold (Chapter 2, (89)). Although the binding affinity for pre-tRNA substrates varies by up to 100-fold, all three *At*PRORPs have similar affinities for a given substrate (Chapter 2, (89)).

Recent biochemical studies also demonstrate that *At*PRORPs do not make contacts with the 3’ trailer or beyond N_{-1} or N_{-2} of the leader that significantly enhance affinity or cleavage rates, therefore recognition determinants are located primarily in the tRNA body (Chapters 2, 3; (88, 89)). Deletion of the D- and anticodon-stems or mutation of the conserved G_{18} nucleotide in the *A*. *thaliana* pre-$tRNA^{Cys}$ abolishes PRORP1-catalyzed cleavage under multiple-turnover conditions (115). However, deletion or elongation of the D-stem in the

cyanobacterial pre-tRNA^{Gln} substrate did not affect cleavage by partially purified *A*. *thaliana* RNase P under multiple turnover conditions (112). These data argue that the D-loop contribution for PRORP recognition may be dependent on the substrate context or stability.

As might be expected, the PPR domain is important for substrate recognition by *At*PRORP1. Deletion of all or part of the PPR repeats decreases substrate affinity by up to 30-fold and abolishes catalysis (37, 55). Conserved nucleotides in D- and TψC-loops (G_{18}, G_{19}, C_{56}, and C_{57}) of mitochondrial pre-tRNA^{Cys} were protected by *At*PRORP1 in footprinting experiments, leading to the hypothesis that the PRORP1 PPR domain contacts this region (115). Other PPR proteins have been implicated in single-stranded RNA (ssRNA) binding with proposed base-specific recognition motifs utilizing residues on two subsequent helices (116, 117). Docking and mutagenesis experiments with *At*PRORP1 suggest that Asn 136 in PPR2 and Thr 180 in PPR3 are involved in substrate recognition (55). Additional mutagenesis studies implicated *At*PRORP3 Arg 145 (*At*PRORP1 Arg 212) in PPR4 in substrate recognition (88). However, contacts between these amino acids and the pre-tRNA substrate are not yet clear (88). Given that many of the D- and TψC-loop bases are typically positioned facing into the structure, making hydrogen bonding and/or stacking interactions, and that PRORPs must recognize all tRNAs, it is likely that PRORP PPRs use a recognition strategy alternative to the ssRNA-binding PPR proteins.

*At*PRORP cleavage site selection is less robust than for bacterial RNase P *in vitro*, resulting in multiple products for various pre-tRNAs (Chapter 2; (84, 89, 118)). The data suggest that this reduced cleavage fidelity results from the ability of PRORPs to recognize an acceptor stem extended by an N_{-1}–N_{73} base pair (Chapter 2; (88, 89)). Reduction of the base-pairing potential or removing the discriminator base restores cleavage at the canonical site (Chapter 2; (88, 89)). Furthermore, *At*PRORPs catalyze reprocessing of miscleaved *A*. *thaliana* plastid pre-tRNA^{Phe} to yield the correct mature tRNA (Chapter 2, (89)). Interestingly, extending the TψC-stem length resulted in a +2 cleavage site, yet reducing the

acceptor stem by an equivalent number of base pairs restored cleavage to the correct site (88). *At*PRORP-catalyzed cleavage of mitochondrial and *E*. *coli* pre-tRNA^{His} occurs mostly at the N_{-2}/N_{-1} phosphodiester bond, thus primarily producing the correct mature tRNA^{His} (88, 118).

Non-tRNA RNase P Substrates

In addition to pre-tRNA, RNase Ps catalyze processing of various other substrates in cells (see (119) for review). Several alternative substrates have been identified for PRORPs, which include a tRNA-like structure (t-element) in a precursor *Arabidopsis* mitochondrial transcript and mirror tRNAs in polycistronic transcripts from the antisense strand of metazoan mitochondrial tRNAs (34, 120). In many, but not all, cases of non-tRNA RNase P substrates, the transcripts are proposed to adopt a structure similar to the tRNA elements that are critical for recognition (121-123).

CONCLUSIONS

RNase P enzymes represent an extreme case of convergent evolution. While RNA-based RNase Ps and eukaryotic PRORPs have distinct evolutionary origins, there are many similarities in their catalytic and substrate recognition strategies. As was detailed by Lechner and colleagues (11), the emergence of PRORP nucleases in early Eukarya appears to have caused a flurry of diversification. The resulting competition for cellular RNase P function produced remarkable diversity, from a lack of PRORP sequences in Fungi (11), to the loss of the RNA-based RNase P in plants such as *Arabidopsis* (43). There remain gaps in our understanding of this class of enzymes, including RNase P *in vivo* substrate selectivity and distribution among several eukaryotic clades.

Some aspects of cleavage activity and substrate recognition mechanisms vary between different classes of RNase P enzymes. Dissecting the mechanism of substrate recognition by RNase Ps will be important for designing inhibitors that specifically target RNase P in pathogens while leaving human RNase P activity intact. Further, mechanistic studies of highly purified eukaryotic RNase P enzymes

are still needed for a deeper understanding of the role of RNase P in processing pre-tRNA *in vitro* and *in vivo*, including the molecular details of substrate recognition by the protein-only human mtRNase P.

In the more than four decades since their initial discovery, much has been learned about the macromolecular composition, catalysis, and substrate recognition of the RNA-based RNase Ps. Likewise, in the nearly eight years since the human mitochondrial PRORP *in vitro* reconstitution was published, much has been learned about the *in vivo* function and domain architecture of the nuclease subunit and its homologs. However, there are still many aspects of catalysis, substrate recognition, and cellular functions that remain to be clarified.

In this thesis, I will present my work, which provides insight into the substrate recognition strategies of the PRORPs from *Arabidopsis thaliana*. We define the structural features of pre-tRNAs that are required for PRORP recognition. We also identify an alternative 5’ end selection behavior by PRORPs *in vitro*. We investigate the molecular determinants of the recognition complex of pre-tRNA substrates and plant PRORPs. We identify a novel tRNA-recognition surface on the PRORP PPR, which is conserved among PRORPs throughout eukaryotes. Finally, we develop a method to map the PRORP-substrate recognition complex.

REFERENCES

1. The Nobel Prize in Chemistry 1989. Available online: http://www.nobelprize.org/nobel_prizes/chemistry/laureates/1989/ (accessed on 15 Janurary 2016).
2. Robertson HD, Altman S, Smith JD. Purification and properties of a Specific *Escherichia coli* ribonuclease which cleaves a tyrosine transfer ribonucleic acid precursor. *J Biol Chem*. 1972;247(16):5243-51.
3. Guerrier-Takada C, Gardiner K, Marsh T, Pace NR, Altman S. The RNA moiety of Ribonuclease P is the catalytic subunit of the enzyme. *Cell*. 1983;35(3 Pt 2):849-57.
4. Frank DN, Pace NR. Ribonuclease P: Unity and diversity in a tRNA processing ribozyme. *Annu Rev Biochem*. 1998;67:153-80.
5. Walker SC, Engelke DR. Ribonuclease P: The evolution of an ancient RNA enzyme. *Crit Rev Biochem Mol Biol*. 2006;41(2):77-102.
6. Hopper AK, Huang HY. Quality control pathways for nucleus-encoded eukaryotic tRNA biosynthesis and subcellular trafficking. *Mol Cell Biol*. 2015;35(12):2052-8.
7. Liu X. Molecular recognition of inhibitors, metal ions and substrates by Ribonuclease P [Ph.D. Thesis]. Ann Arbor, MI: The University of Michigan; 2013.
8. Randau L, Schröder I, Söll D. Life without RNase P. *Nature*. 2008;453(7191):120-3.
9. Randau L. RNA processing in the minimal organism *Nanoarchaeum equitans*. *Genome Biology*. 2012;13(7):1-11.
10. Howard MJ, Liu X, Lim WH, Klemm BP, Fierke CA, Koutmos M, Engelke DR. RNase P enzymes: Divergent scaffolds for a conserved biological reaction. *RNA Biol*. 2013;10(6):909-14.
11. Lechner M, Rossmanith W, Hartmann RK, Thölken C, Gutmann B, Giegé P, Gobert A. Distribution of ribonucleoprotein and protein-Only RNase P in Eukarya. *Mol Biol Evol*. 2015;32(12):3186-93.
12. Adl SM, Simpson AG, Lane CE, Lukes J, Bass D, Bowser SS, Brown MW, Burki F, Dunthorn M, Hampl V, Heiss A, Hoppenrath M, Lara E, Le Gall L, Lynn DH, McManus H, Mitchell EA, Mozley-Stanridge SE, Parfrey LW, Pawlowski J, Rueckert S, Shadwick RS, Schoch CL, Smirnov A, Spiegel FW. The revised classification of eukaryotes. *J Eukaryot Microbiol*. 2012;59(5):429-93.
13. Burki F. The eukaryotic tree of life from a global phylogenomic perspective. *Cold Spring Harb Perspect Biol*. 2014;6(5):a016147.
14. Chen JL, Pace NR. Identification of the universally conserved core of Ribonuclease P RNA. *RNA*. 1997;3(6):557-60.
15. Brown JW, Haas ES, James BD, Hunt DA, Liu JS, Pace NR. Phylogenetic analysis and evolution of RNase P RNA in proteobacteria. *J Bacteriol*. 1991;173(12):3855-63.
16. Haas ES, Brown JW. Evolutionary variation in bacterial RNase P RNAs. *Nucleic Acids Res*. 1998;26(18):4093-9.
17. Brown JW. The Ribonuclease P Database. *Nucleic Acids Res*. 1999;27(1):314.

18. Marquez SM, Harris JK, Kelley ST, Brown JW, Dawson SC, Roberts EC, Pace NR. Structural implications of novel diversity in eucaryal RNase P RNA. *RNA*. 2005;11(5):739-51.
19. Evans D, Marquez SM, Pace NR. RNase P: interface of the RNA and protein worlds. Trends Biochem Sci. 2006;31(6):333-41.
20. Ellis JC, Brown JW. The RNase P family. *RNA Biol*. 2009;6(4):362-9.
21. Chen Y. Ribonuclease P catalysis: Metal ion interaction and conformational dynamics [Ph.D. Thesis]. Ann Arbor, MI: The University of Michigan; 2015.
22. Reiter NJ, Osterman A, Torres-Larios A, Swinger KK, Pan T, Mondragón A. Structure of a bacterial Ribonuclease P holoenzyme in complex with tRNA. *Nature*. 2010;468(7325):784-9.
23. Lim WH. Importance of substrate recognition and metal ions in the Ribonuclease P catalysis [Ph.D. Thesis]. Ann Arbor, MI: The University of Michigan; 2011.
24. Kole R, Altman S. Properties of purified Ribonuclease P from *Escherichia coli*. *Biochemistry*. 1981;20(7):1902-6.
25. Kurz JC, Niranjanakumari S, Fierke CA. Protein component of *Bacillus subtilis* RNase P specifically enhances the affinity for precursor-tRNA^{Asp}. *Biochemistry*. 1998;37(8):2393-400.
26. Kurz JC, Fierke CA. The affinity of magnesium binding sites in the *Bacillus subtilis* RNase P•pre-tRNA complex is enhanced by the protein subunit. *Biochemistry*. 2002;41(30):9545-58.
27. Hsieh J, Andrews AJ, Fierke CA. Roles of protein subunits in RNA-protein complexes: lessons from Ribonuclease P. *Biopolymers*. 2004;73(1):79-89.
28. Hsieh J, Koutmou KS, Rueda D, Koutmos M, Walter NG, Fierke CA. A divalent cation stabilizes the active conformation of the *B*. *subtilis* RNase P•pre-tRNA complex: a role for an inner-sphere metal ion in RNase P. *J Mol Biol*. 2010;400(1):38-51.
29. Koutmou KS, Day-Storms JJ, Fierke CA. The RNR motif of *B*. *subtilis* RNase P protein interacts with both PRNA and pre-tRNA to stabilize an active conformer. *RNA*. 2011;17(7):1225-35.
30. Jarrous N. Human Ribonuclease P: subunits, function, and intranuclear localization. *RNA*. 2002;8(1):1-7.
31. Marvin MC, Engelke DR. Broadening the mission of an RNA enzyme. *J Cell Biochem*. 2009;108(6):1244-51.
32. Holzmann J, Frank P, Löffler E, Bennett KL, Gerner C, Rossmanith W. RNase P without RNA: Identification and functional reconstitution of the human mitochondrial tRNA processing enzyme. *Cell*. 2008;135(3):462-74.
33. Cho IM, Lai LB, Susanti D, Mukhopadhyay B, Gopalan V. Ribosomal protein L7Ae is a subunit of archaeal RNase P. *Proc Natl Acad Sci U S A*. 2010;107(33):14573-8.
34. Gobert A, Gutmann B, Taschner A, Gößringer M, Holzmann J, Hartmann RK, Rossmanith W, Giegé P. A single *Arabidopsis* organellar protein has RNase P activity. *Nat Struct Mol Biol*. 2010;17(6):740-4.

35. Weber C, Hartig A, Hartmann RK, Rossmanith W. Playing RNase P evolution: Swapping the RNA catalyst for a protein reveals functional uniformity of highly divergent enzyme forms. *PLOS Genet*. 2014;10(8):e1004506.
36. Sanchez MI, Mercer TR, Davies SM, Shearwood AM, Nygård KK, Richman TR, Mattick JS, Rackham O, Filipovska A. RNA processing in human mitochondria. *Cell Cycle*. 2011;10(17):2904-16.
37. Howard MJ, Lim WH, Fierke CA, Koutmos M. Mitochondrial Ribonuclease P structure provides insight into the evolution of catalytic strategies for precursor-tRNA 5 ' processing. *Proc Nat Acad Sci U S A*. 2012;109(40):16149-54.
38. Jackman JE, Montange RK, Malik HS, Phizicky EM. Identification of the yeast gene encoding the tRNA m^1G methyltransferase responsible for modification at position 9. *RNA*. 2003;9(5):574-85.
39. Vilardo E, Nachbagauer C, Buzet A, Taschner A, Holzmann J, Rossmanith W. A subcomplex of human mitochondrial RNase P is a bifunctional methyltransferase–extensive moonlighting in mitochondrial tRNA biogenesis. *Nucleic Acids Res*. 2012;40(22):11583-93.
40. Reinhard L, Sridhara S, Hällberg BM. Structure of the nuclease subunit of human mitochondrial RNase P. *Nucleic Acids Res*. 2015;43(11):5664-72.
41. Lai LB, Bernal-Bayard P, Mohannath G, Lai SM, Gopalan V, Vioque A. A functional RNase P protein subunit of bacterial origin in some eukaryotes. *Mol Genet Genomics*. 2011;286(5-6):359-69.
42. Taschner A, Weber C, Buzet A, Hartmann RK, Hartig A, Rossmanith W. Nuclear RNase P of *Trypanosoma brucei*: A single protein in place of the multicomponent RNA-protein complex. *Cell Rep*. 2012;2(1):19-25.
43. Gutmann B, Gobert A, Giegé P. PRORP proteins support RNase P activity in both organelles and the nucleus in *Arabidopsis*. *Genes Dev*. 2012;26(10):1022-7.
44. Green CJ, Rivera-León R, Vold BS. The catalytic core of RNase P. *Nucleic Acids Res*. 1996;24(8):1497-503.
45. Loria A, Pan T. Modular construction for function of a ribonucleoprotein enzyme: The catalytic domain of *Bacillus subtilis* RNase P complexed with *B. subtilis* RNase P protein. *Nucleic Acids Res*. 2001;29(9):1892-7.
46. Wu S, Kikovska E, Lindell M, Kirsebom LA. Cleavage mediated by the catalytic domain of bacterial RNase P RNA. *J Mol Biol*. 2012;422(2):204-14.
47. Harris ME, Nolan JM, Malhotra A, Brown JW, Harvey SC, Pace NR. Use of photoaffinity crosslinking and molecular modeling to analyze the global architecture of Ribonuclease P RNA. *EMBO J*. 1994;13(17):3953-63.
48. Westhof E, Altman S. Three-dimensional working model of M1 RNA, the catalytic RNA subunit of Ribonuclease P from *Escherichia coli*. *Proc Natl Acad Sci U S A*. 1994;91(11):5133-7.
49. Pan T, Loria A, Zhong K. Probing of tertiary interactions in RNA: 2'-hydroxyl-base contacts between the RNase P RNA and pre-tRNA. *Proc Natl Acad Sci U S A*. 1995;92(26):12510-4.
50. Karasik A, Shanmuganathan A, Howard MJ, Fierke CA, Koutmos M. Nuclear Protein-Only Ribonuclease P2 Structure and Biochemical

Characterization Provide Insight into the Conserved Properties of tRNA 5' End Processing Enzymes. *J Mol Biol*. 2016;428(1):26-40.
51. Li F, Liu X, Zhou W, Yang X, Shen Y. Auto-inhibitory Mechanism of the Human Mitochondrial RNase P Protein Complex. *Sci Rep*. 2015;5:9878.
52. Kissinger CR, Rejto PA, Pelletier LA, Thomson JA, Showalter RE, Abreo MA, Agree CS, Margosiak S, Meng JJ, Aust RM, Vanderpool D, Li B, Tempczyk-Russell A, Villafranca JE. Crystal structure of human ABAD/HSD10 with a bound inhibitor: implications for design of Alzheimer's disease therapeutics. *J Mol Biol*. 2004;342(3):943-52.
53. Van Laer B, Roovers M, Wauters L, Kasprzak JM, Dyzma M, Deyaert E, Kumar Singh R, Feller A, Bujnicki JM, Droogmans L, Versées W. Structural and functional insights into tRNA binding and adenosine N1-methylation by an archaeal Trm10 homologue. *Nucleic Acids Res*. 2016;44(2):940-53.
54. Howard MJ, Klemm BP, Fierke CA. Mechanistic studies reveal similar catalytic strategies for phosphodiester bond hydrolysis by protein-only and RNA-dependent Ribonuclease P. *J Biol Chem*. 2015;290(21):13454-64.
55. Imai T, Nakamura T, Maeda T, Nakayama K, Gao X, Nakashima T, Kakuta Y, Kimura M. Pentatricopeptide repeat motifs in the processing enzyme PRORP1 in *Arabidopsis thaliana* play a crucial role in recognition of nucleotide bases at TψC loop in precursor tRNAs. *Biochem Biophys Res Commun*. 2014;450(4):1541-6.
56. Hsieh J, Fierke CA. Conformational change in the *Bacillus subtilis* RNase P holoenzyme–pre-tRNA complex enhances substrate affinity and limits cleavage rate. *RNA*. 2009;15(8):1565-77.
57. Beebe JA, Fierke CA. A kinetic mechanism for cleavage of precursor $tRNA^{Asp}$ catalyzed by the RNA component of *Bacillus subtilis* Ribonuclease P. *Biochemistry*. 1994;33(34):10294-304.
58. Stanford NP, Halford SE, Baldwin GS. DNA cleavage by the EcoRV restriction endonuclease: pH dependence and proton transfers in catalysis. *J Mol Biol*. 1999;288(1):105-16.
59. Bastock JA, Webb M, Grasby JA. The pH-dependence of the *Escherichia coli* RNase HII-catalysed reaction suggests that an active site carboxylate group participates directly in catalysis. *J Mol Biol*. 2007;368(2):421-33.
60. Crary SM, Niranjanakumari S, Fierke CA. The protein component of *Bacillus subtilis* Ribonuclease P increases catalytic efficiency by enhancing interactions with the 5' leader sequence of pre-$tRNA^{Asp}$. *Biochemistry*. 1998;37(26):9409-16.
61. Loria A, Pan T. Recognition of the 5' leader and the acceptor stem of a pre-tRNA substrate by the ribozyme from *Bacillus subtilis* RNase P. *Biochemistry*. 1998;37(28):10126-33.
62. Loria A, Pan T. The cleavage step of Ribonuclease P catalysis is determined by ribozyme-substrate interactions both distal and proximal to the cleavage site. *Biochemistry*. 1999;38(27):8612-20.
63. Brännvall M, Kikovska E, Wu S, Kirsebom LA. Evidence for induced fit in bacterial RNase P RNA-mediated cleavage. *J Mol Biol*. 2007;372(5):1149-64.
64. Misra VK, Draper DE. On the role of magnesium ions in RNA stability. *Biopolymers*. 1998;48(2-3):113-35.

65. Smith D, Pace NR. Multiple magnesium ions in the Ribonuclease P reaction mechanism. *Biochemistry*. 1993;32(20):5273-81.
66. Pan T. Higher order folding and domain analysis of the ribozyme from *Bacillus subtilis* Ribonuclease P. *Biochemistry*. 1995;34(3):902-9.
67. Beebe JA, Kurz JC, Fierke CA. Magnesium ions are required by *Bacillus subtilis* Ribonuclease P RNA for both binding and cleaving precursor $tRNA^{Asp}$. *Biochemistry*. 1996;35(32):10493-505.
68. Christian EL, Smith KM, Perera N, Harris ME. The P4 metal binding site in RNase P RNA affects active site metal affinity through substrate positioning. *RNA*. 2006;12(8):1463-7.
69. Harris ME, Pace NR. Identification of phosphates involved in catalysis by the ribozyme RNase P RNA. *RNA*. 1995;1(2):210-8.
70. Kazantsev AV, Pace NR. Identification by modification-interference of purine N-7 and ribose 2'-OH groups critical for catalysis by bacterial Ribonuclease P. *RNA*. 1998;4(8):937-47.
71. Crary SM, Kurz JC, Fierke CA. Specific phosphorothioate substitutions probe the active site of *Bacillus subtilis* Ribonuclease P. *RNA*. 2002;8(7):933-47.
72. Christian EL, Kaye NM, Harris ME. Evidence for a polynuclear metal ion binding site in the catalytic domain of Ribonuclease P RNA. *EMBO J*. 2002;21(9):2253-62.
73. Warnecke JM, Fürste JP, Hardt WD, Erdmann VA, Hartmann RK. Ribonuclease P (RNase P) RNA is converted to a Cd^{2+}-ribozyme by a single R_P-phosphorothioate modification in the precursor tRNA at the RNase P cleavage site. *Proc Natl Acad Sci U S A*. 1996;93(17):8924-8.
74. Chen Y, Li X, Gegenheimer P. Ribonuclease P catalysis requires Mg^{2+} coordinated to the pro-R_P oxygen of the scissile bond. *Biochemistry*. 1997;36(9):2425-38.
75. Warnecke JM, Held R, Busch S, Hartmann RK. Role of metal ions in the hydrolysis reaction catalyzed by RNase P RNA from *Bacillus subtilis*. *J Mol Biol*. 1999;290(2):433-45.
76. Pfeiffer T, Tekos A, Warnecke JM, Drainas D, Engelke DR, Seraphin B, Hartmann RK. Effects of phosphorothioate modifications on precursor tRNA processing by eukaryotic RNase P enzymes. *J Mol Biol*. 2000;298(4):559-65.
77. Persson T, Cuzic S, Hartmann RK. Catalysis by RNase P RNA: unique features and unprecedented active site plasticity. *J Biol Chem*. 2003;278(44):43394-401.
78. Cassano AG, Anderson VE, Harris ME. Analysis of solvent nucleophile isotope effects: evidence for concerted mechanisms and nucleophilic activation by metal coordination in nonenzymatic and ribozyme-catalyzed phosphodiester hydrolysis. *Biochemistry*. 2004;43(32):10547-59.
79. Cassano AG, Anderson VE, Harris ME. Understanding the transition states of phosphodiester bond cleavage: insights from heavy atom isotope effects. *Biopolymers*. 2004;73(1):110-29.
80. Steitz TA, Steitz JA. A general two-metal-ion mechanism for catalytic RNA. *Proc Natl Acad Sci U S A*. 1993;90(14):6498-502.

81. Brautigam CA, Sun S, Piccirilli JA, Steitz TA. Structures of normal single-stranded DNA and deoxyribo-3'-S-phosphorothiolates bound to the 3'-5' exonucleolytic active site of DNA polymerase I from *Escherichia coli*. *Biochemistry*. 1999;38(2):696-704.
82. Sosunov V, Sosunova E, Mustaev A, Bass I, Nikiforov V, Goldfarb A. Unified two-metal mechanism of RNA synthesis and degradation by RNA polymerase. *EMBO J*. 2003;22(9):2234-44.
83. Sun L, Harris ME. Evidence that binding of C5 protein to P RNA enhances ribozyme catalysis by influencing active site metal ion affinity. *RNA*. 2007;13(9):1505-15.
84. Pavlova LV, Gößringer M, Weber C, Buzet A, Rossmanith W, Hartmann RK. tRNA processing by protein-only versus RNA-based RNase P: kinetic analysis reveals mechanistic differences. *ChemBioChem*. 2012;13(15):2270-6.
85. Thomas BC, Li X, Gegenheimer P. Chloroplast ribonuclease P does not utilize the ribozyme-type pre-tRNA cleavage mechanism. *RNA*. 2000;6(4):545-53.
86. Westhof E, Dumas P, Moras D. Restrained refinement of two crystalline forms of yeast aspartic acid and phenylalanine transfer RNA crystals. *Acta Crystallogr A*. 1988;44 (Pt 2):112-23.
87. Orans J, McSweeney EA, Iyer RR, Hast MA, Hellinga HW, Modrich P, Beese LS. Structures of human exonuclease 1 DNA complexes suggest a unified mechanism for nuclease family. *Cell*. 2011;145(2):212-23.
88. Brillante N, Gößringer M, Lindenhofer D, Toth U, Rossmanith W, Hartmann RK. Substrate recognition and cleavage-site selection by a single-subunit protein-only RNase P. *Nucleic Acids Res*. 2016;44(5):2323-36.
89. Howard MJ, Karasik A, Klemm BP, Mei C, Shanmuganathan A, Fierke CA, Koutmos M. Differential substrate recognition by isozymes of plant protein-only Ribonuclease P. *RNA*. 2016;22(5):782-92.
90. Hsieh J, Walker SC, Fierke CA, Engelke DR. Pre-tRNA turnover catalyzed by the yeast nuclear RNase P holoenzyme is limited by product release. *RNA*. 2009;15(2):224-34.
91. Dirheimer G, Keith G, Dumas P, Westhof E. Primary, Secondary, and Tertiary Structures of tRNAs. In: *tRNA: Structure, Biosynthesis, and Function*, 1st ed.; Söll, D., RajBhandary, U.L., Eds.; The American Society for Microbiology Press: Washington, DC, USA, 1995; Chapter 8; pp. 93-126.
92. Suzuki T, Nagao A, Suzuki T. Human mitochondrial tRNAs: biogenesis, function, structural aspects, and diseases. *Annu Rev Genet*. 2011;45:299-329.
93. Rossmanith W. Of P and Z: mitochondrial tRNA processing enzymes. *Biochim Biophys Acta*. 2012;1819(9-10):1017-26.
94. Byrne RT, Konevega AL, Rodnina MV, Antson AA. The crystal structure of unmodified $tRNA^{Phe}$ from *Escherichia coli*. *Nucleic Acids Res*. 2010;38:4154-62.
95. Rueda D, Hsieh J, Day-Storms JJ, Fierke CA, Walter NG. The 5' leader of precursor $tRNA^{Asp}$ bound to the *Bacillus subtilis* RNase P holoenzyme has an extended conformation. *Biochemistry*. 2005;44(49):16130-9.
96. Sun L, Campbell FE, Zahler NH, Harris ME. Evidence that substrate-specific effects of C5 protein lead to uniformity in binding and catalysis by RNase P. *Embo J*. 2006;25(17):3998-4007.

97. Niranjanakumari S, Stams T, Crary SM, Christianson DW, Fierke CA. Protein component of the ribozyme ribonuclease P alters substrate recognition by directly contacting precursor tRNA. *Proc Natl Acad Sci U S A*. 1998;95(26):15212-7.
98. Tallsjo A, Kirsebom LA. Product release is a rate-limiting step during cleavage by the catalytic RNA subunit of *Escherichia coli* RNase P. *Nucleic Acids Res*. 1993;21(1):51-7.
99. Oh BK, Pace NR. Interaction of the 3'-end of tRNA with ribonuclease P RNA. *Nucleic Acids Res*. 1994;22(20):4087-94.
100. Christian EL, Zahler NH, Kaye NM, Harris ME. Analysis of substrate recognition by the ribonucleoprotein endonuclease RNase P. *Methods*. 2002;28(3):307-22.
101. Zahler NH, Christian EL, Harris ME. Recognition of the 5' leader of pre-tRNA substrates by the active site of ribonuclease P. *RNA*. 2003;9(6):734-45.
102. Wegscheid B, Hartmann RK. The precursor tRNA 3'-CCA interaction with *Escherichia coli* RNase P RNA is essential for catalysis by RNase P *in vivo*. *RNA*. 2006;12(12):2135-48.
103. Wegscheid B, Hartmann RK. *In vivo* and *in vitro* investigation of bacterial type B RNase P interaction with tRNA 3'-CCA. *Nucleic Acids Res*. 2007;35(6):2060-73.
104. Koutmou KS, Zahler NH, Kurz JC, Campbell FE, Harris ME, Fierke CA. Protein-Precursor tRNA Contact Leads to Sequence-Specific Recognition of 5' Leaders by Bacterial Ribonuclease P. *J Mol Biol*. 2010;396(1):195-208.
105. Kirsebom LA, Svard SG. Base pairing between *Escherichia coli* RNase P RNA and its substrate. *Embo J*. 1994;13(20):4870-6.
106. Oh BK, Frank DN, Pace NR. Participation of the 3'-CCA of tRNA in the binding of catalytic Mg^{2+} ions by ribonuclease P. *Biochemistry*. 1998;37(20):7277-83.
107. Heide C, Pfeiffer T, Nolan JM, Hartmann RK. Guanosine 2-NH2 groups of *Escherichia coli* RNase P RNA involved in intramolecular tertiary contacts and direct interactions with tRNA. *RNA*. 1999;5(1):102-16.
108. Busch S, Kirsebom LA, Notbohm H, Hartmann RK. Differential role of the intermolecular base-pairs G292-C(75) and G293-C(74) in the reaction catalyzed by *Escherichia coli* RNase P RNA. *J Mol Biol*. 2000;299(4):941-51.
109. Heide C, Busch S, Feltens R, Hartmann RK. Distinct modes of mature and precursor tRNA binding to *Escherichia coli* RNase P RNA revealed by NAIM analyses. *RNA*. 2001;7(4):553-64.
110. Zahler NH, Sun L, Christian EL, Harris ME. The pre-tRNA nucleotide base and 2'-hydroxyl at N(-1) contribute to fidelity in tRNA processing by RNase P. *J Mol Biol*. 2005;345(5):969-85.
111. Brännvall M, Mattsson JG, Svard SG, Kirsebom LA. RNase P RNA structure and cleavage reflect the primary structure of tRNA genes. *J Mol Biol*. 1998;283(4):771-83.
112. Chen WY, Singh D, Lai LB, Stiffler MA, Lai HD, Foster MP, Gopalan V. Fidelity of tRNA 5'-maturation: a possible basis for the functional dependence of

archaeal and eukaryal RNase P on multiple protein cofactors. *Nucleic Acids Res*. 2012;40(10):4666-80.
113. Maréchal-Drouard L, J H Weil a, Dietrich A. Transfer RNAs and Transfer RNA Genes in Plants. *Annu Rev Plant Physiol Plant Mol Biol*. 1993;44(1):13-32.
114. Cognat V, Pawlak G, Duchêne AM, Daujat M, Gigant A, Salinas T, Michaud M, Gutmann B, Giegé P, Gobert A, Maréchal-Drouard L. PlantRNA, a database for tRNAs of photosynthetic eukaryotes. *Nucleic Acids Res*. 2013;41(Database issue):D273-9.
115. Gobert A, Pinker F, Fuchsbauer O, Gutmann B, Boutin R, Roblin P, Sauter C, Giegé P. Structural insights into protein-only RNase P complexed with tRNA. *Nat Commun*. 2013;4(1353):1-8.
116. Barkan A, Rojas M, Fujii S, Yap A, Chong YS, Bond CS, Small I. A Combinatorial Amino Acid Code for RNA Recognition by Pentatricopeptide Repeat Proteins. *PLOS Genetics*. 2012;8(8):e1002910.
117. Yagi Y, Hayashi S, Kobayashi K, Hirayama T, Nakamura T. Elucidation of the RNA recognition code for pentatricopeptide repeat proteins involved in organelle RNA editing in plants. *PLOS One*. 2013;8(3):e57286.
118. Placido A, Sieber F, Gobert A, Gallerani R, Giegé P, Maréchal-Drouard L. Plant mitochondria use two pathways for the biogenesis of $tRNA^{His}$. *Nucleic Acids Res*. 2010;38(21):7711-7.
119. Hernandez-Cid A, Aguirre-Sampieri S, Diaz-Vilchis A, Torres-Larios A. Ribonucleases P/MRP and the expanding ribonucleoprotein world. *IUBMB Life*. 2012;64(6):521-8.
120. Rackham O, Shearwood AM, Mercer TR, Davies SM, Mattick JS, Filipovska A. Long noncoding RNAs are generated from the mitochondrial genome and regulated by nuclear-encoded proteins. *RNA*. 2011;17(12):2085-93.
121. Komine Y, Kitabatake M, Yokogawa T, Nishikawa K, Inokuchi H. A tRNA-like structure is present in 10Sa RNA, a small stable RNA from *Escherichia coli*. *Proc Natl Acad Sci U S A*. 1994;91(20):9223-7.
122. Wilusz JE, Freier SM, Spector DL. 3' End Processing of a Long Nuclear-Retained Noncoding RNA Yields a tRNA-like Cytoplasmic RNA. *Cell*. 2008;135(5):919-32.
123. Sunwoo H, Dinger ME, Wilusz JE, Amaral PP, Mattick JS, Spector DL. MEN epsilon/beta nuclear-retained non-coding RNAs are up-regulated upon muscle differentiation and are essential components of paraspeckles. *Genome Res*. 2009;19(3):347-59.

CHAPTER 2

Differential substrate recognition by isozymes of plant protein-only Ribonuclease P†

ABSTRACT

Ribonuclease P (RNase P) catalyzes the cleavage of leader sequences from precursor tRNA (pre-tRNA). Typically, these enzymes are ribonucleic protein complexes that are found in all domains of life. However, a new class of RNase P has been discovered that is composed entirely of protein, termed protein-only RNase P (PRORP). To investigate the molecular determinants of PRORP substrate recognition, we measured the binding affinities and cleavage kinetics of *Arabidopsis* PRORP1 for varied pre-tRNA substrates. This analysis revealed that PRORP1 does not make significant contacts within the trailer or beyond N_{-1} of the leader, indicating that this enzyme recognizes primarily the tRNA body. To determine the extent to which sequence variation within the tRNA body modulates substrate selectivity and to provide insight into the evolution and function of PRORP enzymes, we measured the reactivity of the three *Arabidopsis* PRORP isozymes (PRORP1–3) with four pre-tRNA substrates. A 13-fold range in catalytic efficiencies (10^4-10^5 M^{-1} s^{-1}) was observed, demonstrating moderate selectivity for

†This chapter is reformatted from reference: Howard, MJ; Karasik, A; Klemm, BP; Mei, C; Shanmuganathan, A; Fierke, CA; Koutmos, M. Differential substrate recognition by isozymes of plant protein-only Ribonuclease P. *RNA* 2016, 22(5):782-92.

M.J.H. prepared the substrates, performed MTO, STO, and binding assays with PRORP1, analyzed the data, made the figures, and wrote a draft of the paper. M.J.H. and B.P.K. designed experiments. B.P.K. edited the paper, prepared PRORP3 constructs, performed assays with PRORP3, and responded to the reviewers, including designing and performing the primer extension assays and analyzing these data. A.K. performed assays with PRORP2. C.M. purified PRORP3 and performed assays with model substrates. A.S. purified PRORP2.

pre-tRNA substrates. Although PRORPs bind the different pre-tRNA species with affinities varying by as much as 100-fold, the three isozymes have similar affinities for a given pre-tRNA, suggesting similar binding modes. However, PRORP isozymes have varying degrees of cleavage fidelity, which is dependent on the pre-tRNA species and the presence of a 3′-discriminator base. This work defines molecular determinants of PRORP substrate recognition that provides insight into this new class of RNA processing enzymes.

INTRODUCTION

Transfer RNAs (tRNAs) are transcribed as precursors (pre-tRNA) that contain extra nucleotides flanking the 5′ and 3′ ends. Removal of these extraneous sequences is critical for tRNA function, thus the enzymes responsible for tRNA maturation are essential. RNase P is the endonuclease responsible for catalyzing the 5′ end maturation of pre-tRNA and is found in all domains of life (1). RNase P enzymes are extremely diverse with regard to their macromolecular composition (2). In Bacteria, Archaea, and some eukaryotic lineages, pre-tRNA cleavage is catalyzed by a conserved RNA-based RNase P (ribozyme) that associates with a number of different protein cofactors (1–10, depending on the domain) (3). However, the majority of eukaryotic lineages are predicted to use a protein-only form of RNase P (PRORP) to catalyze pre-tRNA maturation within organelles (mitochondria and chloroplast) and, depending on the lineage, within the nucleus (3, 4). PRORPs were first discovered in human mitochondria, where they process mitochondrial encoded pre-tRNAs (5). Human mitochondrial RNase P requires three protein subunits for efficient catalysis: a nuclease (MRPP3/human PRORP), a tRNA methyltransferase (TRMT10C/MRPP1), and a dehydrogenase (SDR5C1/MRPP2). SDR5C1 and TRMT10C form a complex and are proposed to play a scaffolding role in pre-tRNA maturation catalyzed by human PRORP (6, 7). In contrast to the mammalian enzyme, the recombinant PRORPs from plants, some protists, and algae do not require additional proteins for efficient catalysis, offering a simpler model system to understand the function of this new class of nuclease (4, 8, 9).

While mammals retain an RNA-based RNase P in their nucleus, bioinformatic studies suggested that land plants lack a catalytic RNA component for RNase P activity (10). Consistent with this, the prototypical land plant *Arabidopsis thaliana* possesses three nuclear-encoded PRORP enzymes (PRORP1, 2, and 3; Fig. 2-1A) (11, 12). PRORP1 is essential and is localized to the mitochondria and chloroplasts (4, 12). PRORP2 and 3 are localized to the nucleus where they play essential but redundant roles; knockout of both PRORP2 and PRORP3 is lethal; however, knockout of either one is not (4, 12, 13). Similar to *Arabidopsis*, the moss *Physcomitrella patens* contains three PRORP isozymes and no apparent catalytic RNA component for the RNA-dependent RNase P (14). In contrast to *Arabidopsis*, *P*. *patens* localizes two PRORP enzymes to the mitochondria and chloroplast and one to the nucleus (14). Thus, PRORP enzymes may have replaced the ancient RNA-based RNase P in *A*. *thaliana* and *P*. *patens*, processing tRNA transcripts in the chloroplast, mitochondria, and nucleus. However, the extent to which these isozymes vary in substrate specificity and the molecular interactions that confer substrate specificity of these important enzymes remains largely unknown.

PRORP enzymes contain three domains: an N-terminal pentatricopeptide repeat (PPR) domain, a central Zn-binding domain, and a Nedd4-BP1, YacP nuclease (NYN) domain (Fig. 2-1) (15). The PPR domain contains five PPR or PPR-like motifs, which are helix–turn–helix motifs found in tandem. This domain has been proposed to interact with pre-tRNA and enhance binding affinity (15-17). The largest sequence variation among the PRORPs is found in the PPR domains and thus differences in substrate recognition may lie within this region. The central domain binds a Zn^{2+} ion and structurally links the PPR and metallonuclease domains. The NYN-metallonuclease domain catalyzes phosphodiester bond hydrolysis and contains four conserved aspartates important for binding catalytic metal ions and catalysis (15, 18). Both the protein and RNA-based RNase P enzymes are proposed to use a two-metal ion mechanism for catalysis (18, 19). Despite this similarity, RNA-based RNase P enzymes have higher catalytic efficiency under *in vitro* conditions than PRORP enzymes (2). The sequence identity among the *A*. *thaliana* PRORPs is highest in the metallonuclease domain.

PRORP2 and 3 are most similar (80% identity and 88% similarity); the percent identity and similarity between PRORP1 and 2 are 48% and 65% and between PRORP1 and 3 are 49% and 65%, respectively.

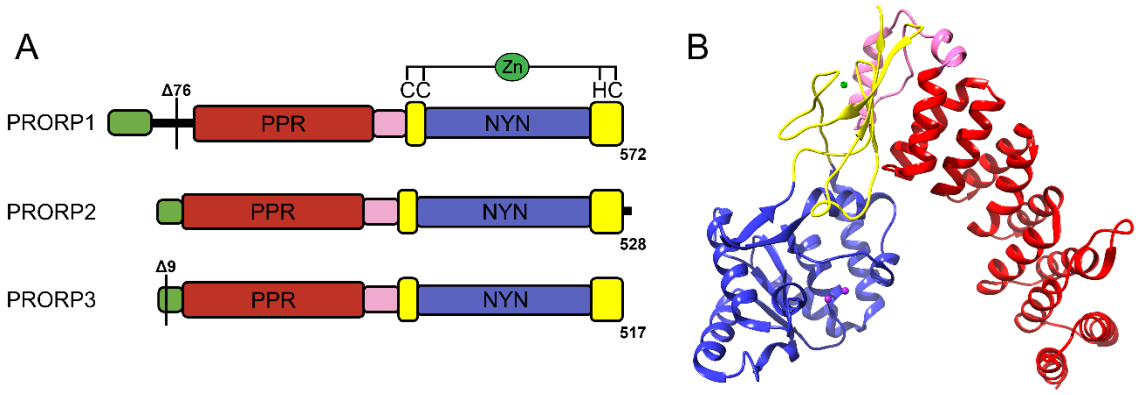


Figure 2-1: (**A**) *A. thaliana* encodes three PRORP enzymes, PRORP1, 2, and 3. N-terminal truncations of PRORP1 (Δ76) and PRORP3 (Δ9) were used in this study. PRORP enzymes contain three domains: PPR (red), central-Zn binding (yellow), and NYN (blue) domains. The N-terminal green region represents the proposed localization sequences: PRORP1 contains a mitochondrial targeting peptide (1–48, predicted by the TargetP 1.1 algorithm (20)), while PRORP2 and 3 contain canonical bipartite nuclear localization signals (H/R-RSR-R/H-X9-K-K-K-K) (21). The pink region represents a plant-specific insert in the gene sequence. (**B**) The crystal structure of PRORP1 (PDB 4G24) with the domains colored as in A.

To provide insight into the evolution and function of PRORP enzymes, we measured substrate specificities, equilibrium binding affinities, and cleavage fidelities for varied pre-tRNA substrates. Previous studies have identified the importance of the elbow region of tRNA for PRORP1 substrate recognition (16, 17). However, recognition of the tRNA leader and trailer by PRORP is largely unknown. Our data demonstrate no dependence on the trailer or leader length beyond the first nucleotide (N_{-1}). This is in contrast to bacterial RNA-based RNase P that makes significant interactions with the leader and trailer sequences of pre-tRNA (22-25). However, the pre-tRNA sequence alters the binding affinity (≤ 40-fold) and, hence, the catalytic efficiency of PRORP cleavage. The three PRORP isozymes have comparable catalytic efficiencies for a given pre-tRNA suggesting similar, but not identical, substrate selectivity. However, PRORP isozymes have varying degrees of cleavage fidelity, which is dependent on the pre-tRNA species and the presence of a 3'-discriminator base. This work defines molecular determinants of PRORP substrate recognition that provides insight into this new class of RNA processing enzymes.

RESULTS

PRORP1 substrate recognition has little dependence on leader and trailer length

The binding affinity of *Bacillus subtilis* RNase P for pre-tRNA has a significant dependence on pre-tRNA leader length, with the pre-tRNA affinity increasing ~ 50-fold when the leader length is increased from one to five nucleotides (26). To determine whether PRORP1 shares a similar dependence on leader length, we measured the binding affinity and single-turnover (STO) cleavage rate constant for *B*. *subtilis* pre-tRNA^{Asp} lacking the 3′-CCA sequence. This substrate was previously used to interrogate both the mechanism of PRORP (18) and leader interactions with *B*. *subtilis* RNase P (22, 26, 27). This allows a direct comparison between the recognition modes of PRORP1 and bacterial RNase P.

A fluorescence anisotropy binding assay was used to measure the equilibrium dissociation constant (K_D) of PRORP1 binding to 5′ fluorescein-labeled *B*. *subtilis* pre-tRNAs with varying leader lengths while maintaining a discriminator base on the 3′ end (Fig. 2-2A). Increasing the pre-tRNA leader length from 1- to 14-nt has little effect on PRORP1 binding affinity (Table 2-1; Fig. 2-2B), with K_D values ranging between 100 and 310 nM. The observed binding trend is independent of the presence of the 5′-fluorescein label on the 5′ end of pre-tRNA (Fig. A-1). In contrast, the binding affinity of the mature tRNA product (0-nt leader) is ~ 30-fold weaker than for pre-tRNA. These data suggest that PRORP1 does not significantly recognize the leader past the first nucleotide (N_{-1}).

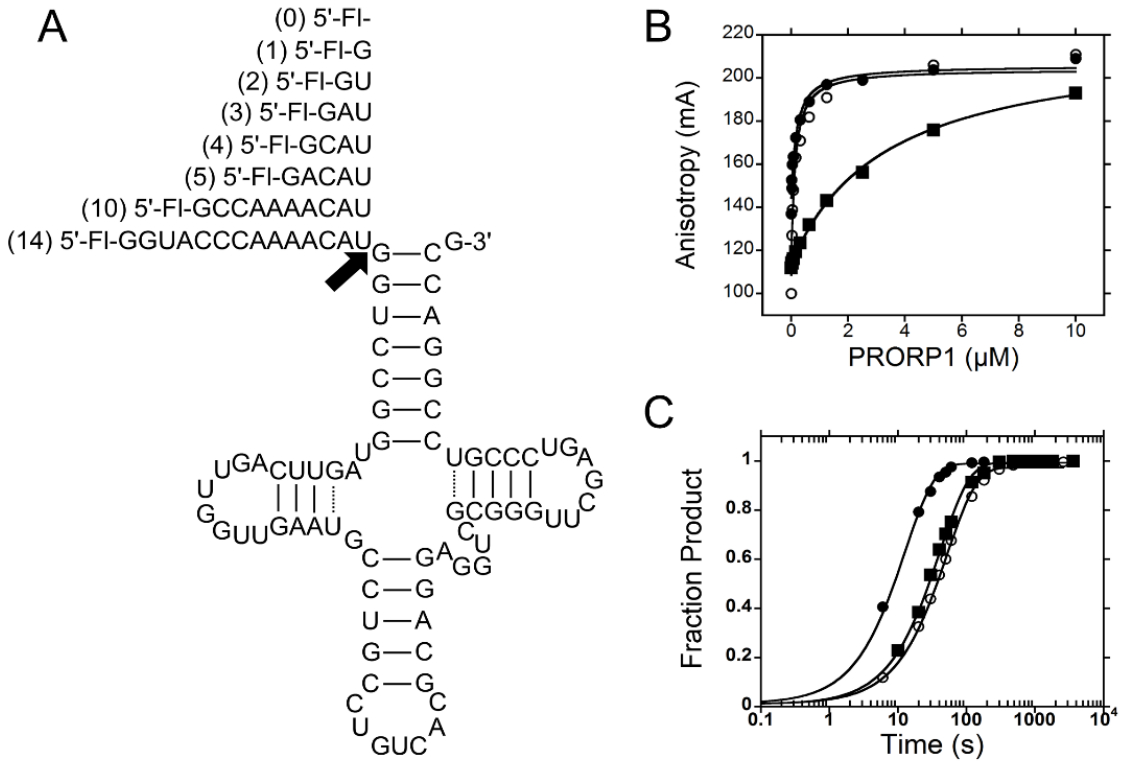


Figure 2-2: Binding affinity and STO activity of PRORP1 for varying leader lengths of *B*. *subtilis* pre-tRNA. (**A**) Predicted secondary structures of B. subtilis pre-tRNAs with varying leader lengths used in this study. All substrates were prepared using *in vitro* transcription. The 5′ end is labeled with fluorescein (Fl). The arrow indicates the RNase P cleavage site. (**B**) Representative binding isotherms for 0:1 (mature tRNA with a discriminator base at the 3′ end) (closed square), 1:1 (closed circle), and 14:1 (open circle). Values for STO rate constants and K_D's are reported in Table 2-1. (**C**) Representative STO timecourses for PRORP1 catalyzed *B*. *subtilis* pre-tRNA cleavage under standard reaction conditions; 1:1 (closed circle), 5:1 (closed square), and 14:1 (open circle).

Table 2-1. Dissociation constant (K_D) and STO observed rate constant (k_{obs}) for PRORP1 with *B. subtilis* pre-tRNA containing varying leader lengths.

pre-tRNA	K_D (nM)[a]	Fold [b]	k_{obs} (s^{-1})[c]	Fold [b]
0	3,400 ± 400	34	–	–
1	150 ± 60	1.5	0.078 ± 0.003	3.9
2	310 ± 20	3.1	0.15 ± 0.02	7.8
3	140 ± 40	1.4	0.032 ± 0.001	1.6
4	150 ± 40	1.5	0.025 ± 0.001	1.3
5	190 ± 60	1.9	0.025 ± 0.001	1.3
10	100 ± 50	1	0.025 ± 0.001	1.3
14	100 ± 30	1	0.02 ± 0.001	1

[a] Measured as described in the legend of Fig. 2-2 and Materials and Methods. The mean and standard deviation is reported from two independent determinations.

[b] As compared to pre-tRNA with a 14-nt leader.

[c] Measured as described in the legend of Fig. 2-2 and Materials and Methods. The standard error from fitting is reported.

To examine the effect of leader length on PRORP1 cleavage activity, the STO rate constant (k_{obs}) was measured using a gel-based assay. The observed rate constant under these conditions is proposed to mainly reflect the chemical cleavage step (18). In all cases, the leader length was consistent with cleavage at the correct site. The STO rate constant for *B*. *subtilis* pre-tRNA with varying leader lengths was generally uniform (Table 2-1; Fig. 2-2C), with the exception of the 1- and 2-nt leader substrates, which have approximately four- and eight-fold increased STO rate constant compared to 14:1 (leader nucleotide length: trailer nucleotide length) pre-tRNA, respectively. This increased reactivity may be due to catalytic enhancement imparted by the positioning of the 5′ guanosine and fluorescein. The cleavage rate constant of the 1:1 pre-tRNA may depend on the sequence at both the -1 nt and the discriminatory base, as suggested by the slower STO cleavage rate of an unmodified U-tRNA (Fig. A-1). Nonetheless, these data indicate that PRORP1 can efficiently bind and catalyze removal of 1-nt leader *B*. *subtilis* pre-tRNA, further demonstrating that PRORP1 does not significantly contact the leader sequence beyond N_{-1}.

To determine whether the trailer is an important determinant for PRORP1 recognition, we assayed the STO activity and binding affinity of PRORP1 with *A*. *thaliana* mitochondrial encoded Cys-Mito pre-tRNA (Fig. 2-3) containing a 24-nt long trailer and a 5-nt long leader (5:24). A Cys-Mito pre-tRNA substrate was used for these studies because extension of the *B*. *subtilis* pre-tRNA trailer to 20-nt resulted in significant STO biphasic kinetics (data not shown), potentially as a result of the pre-tRNA adopting an alternative structure (28). Variation of the Cys-Mito trailer length does not alter the binding affinity (K_D = 65 ± 5 and 70 ± 10 nM for 5:1 and 5:24, respectively) or STO cleavage rate constant (k_{obs} = 0.037 ± 0.002 and 0.033 ± 0.001 s^{-1} for 5:1 and 5:24, respectively), suggesting that PRORP1 does not make significant contacts with the 3′ trailer (Table A-1). Taken together, these data suggest that the leader (past N_{-1}) and trailer sequences are not important determinants of molecular recognition by PRORP1, therefore PRORP1 mainly recognizes the tRNA body.

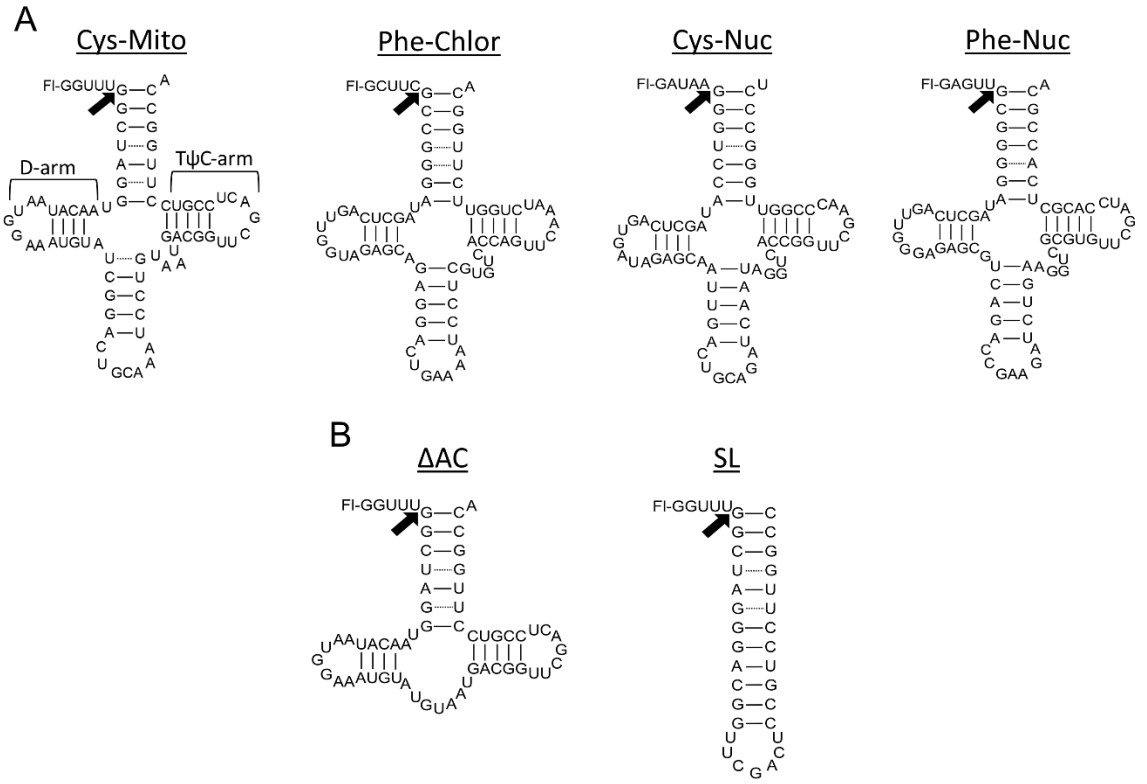


Figure 2-3: (**A**) Predicted secondary structures of *A*. *thaliana* pre-tRNA substrates used in this study. All substrates were prepared using *in vitro* transcription. Cys-Mito and Phe-Chlor substrates are derived from the organellar genomes (Cys-mitochondrial and Phe-chloroplast), and Cys-nuclear and Phe-nuclear are from the nuclear genome. The pre-tRNAs contain 5-nt leaders labeled with fluorescein at the 5′ end and a discriminator base at the 3′ end. (**B**) Model substrates. Predicted secondary structures of two model substrates, ΔAC and SL, derived from the mitochondrial pre-tRNA^{Cys} sequence. The SL substrate connects the acceptor stem helix and TψC-arm.

PRORP reactivity with model substrates reveals the importance of the D-arm in recognition

To determine the regions within the tRNA body that are important for PRORP recognition, we created two truncated versions of pre-tRNA Cys-Mito. These model substrates include an RNA lacking an anti-codon arm (ΔAC), and an RNA that connects the acceptor arm and TψC-arm to form a stem-loop (SL), thereby removing the D- and anticodon arms (Fig. 2-3). Both were labeled at the 5’ end with fluorescein and cleavage catalyzed by PRORP1 was assessed under STO conditions. PRORP1 catalyzed removal of the 5’ leader from the ΔAC substrate with an observed rate constant similar to pre-tRNA, k_{obs} = 0.037 ± 0.002 and 0.023 ± 0.005 s^{-1} for Cys-Mito pre-tRNA and ΔAC, respectively. PRORP1 binding affinity

for the ΔAC and pre-tRNA substrates are similar: K_D = 70 ± 10 and 65 ± 5 nM, respectively (under standard binding conditions with the exception of 100 mM NaCl). This result is consistent with a previous observation that PRORPs can cleave t-element RNA, tRNA-like structures that lack an anticodon stem–loop (4, 12). The SL RNA was cleaved > 1000-fold slower than Cys-Mito pre-tRNA (k_{obs} < 3.3 × 10^{-5} s^{-1}). This result indicates that the D-arm of pre-tRNA is critical for substrate binding and/or cleavage by PRORP1.

Arabidopsis PRORP isozymes display differential catalytic efficiencies with four pre-tRNAs

Given that the PRORP1 substrate recognition relies on interactions with the tRNA body, differences in the nucleotide sequence of tRNA species could potentially alter recognition and catalytic efficiencies. To explore the range of reactivity, we measured the steady-state (multiple-turnover, MTO) kinetic parameters of the three *A*. *thaliana* PRORPs with four pre-tRNAs. The pre-tRNAs assayed are: two nuclear-encoded pre-tRNAs (Cys-Nuc and Phe-Nuc) and two organellar-encoded pre-tRNAs (Cys-Mito and Phe-Chlor) (Fig. 2-3). We used a real-time fluorescence anisotropy assay to measure the steady-state kinetic parameters for PRORP-catalyzed pre-tRNA hydrolysis (Table 2-2; Fig. 2-4A) (29). A comparison of the catalytic efficiencies (k_{cat}/K_M values) reveals that PRORP1 has the highest activity, with values two- to 13-fold higher for a given substrate than PRORP2 and 3 (Fig. 2-4B). The kinetic parameter k_{cat}/K_M is a measure of the productive associations between enzyme and substrate and sets a lower limit for the second-order rate constant for substrate binding (k_{on}). PRORP1 is most selective for the Phe-Chlor substrate (k_{cat}/K_M = 4.1 × 10^5 M^{-1} s^{-1}) relative to PRORP2 and 3, representing a nine- and 13-fold enhancement, respectively. Each isozyme reacts fastest with a different pre-tRNA: PRORP1 with Phe-Chlor, PRORP2 with Phe-Nuc, and PRORP3 with Cys-Mito. The values of the turnover number (k_{cat}) for cleavage catalyzed by PRORP1 and 2 vary little (approximately two-fold) between substrates. However, PRORP3 catalyzes cleavage of the Cys-Nuc substrate nearly eight-fold slower than Phe-Nuc pre-tRNA. These data

indicate that sequence variation within the tRNA body can modulate substrate specificity, albeit within a range of 10^4–10^5 M^{-1} s^{-1}.

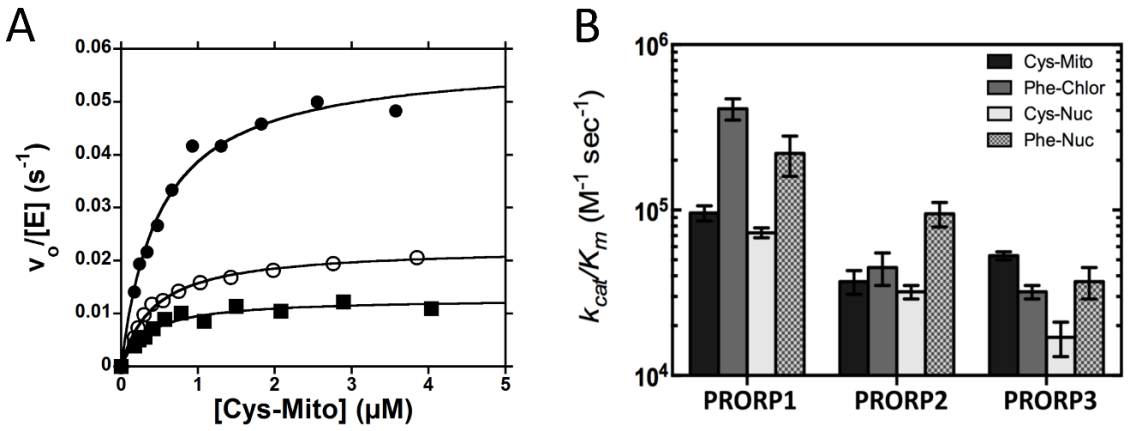


Figure 2-4: Multiple-turnover cleavage of pre-tRNA catalyzed by PRORPs. (**A**) The substrate dependence of the initial velocity (v_0/[E]) for cleavage of Cys-Mito pre-tRNA catalyzed by PRORP1 (closed circle), PRORP2 (closed square), and PRORP3 (open circle). Reactions were performed under standard reaction conditions. (**B**) Bar graph comparing the catalytic efficiencies (k_{cat}/K_M) of the three PRORP enzymes with the four pre-tRNA substrates. The error bars represent the standard error from fitting.

Table 2-2. Kinetic parameters for pre-tRNA cleavage catalyzed by *A. thaliana* PRORP1, 2 and 3.

		Single-turnover[a]	*Multiple-turnover*[b]		
Enzyme	Pre-tRNA	k_{obs} (s^{-1})[c]	k_{cat} (s^{-1})	K_M (nM)	k_{cat}/K_M ($mM^{-1}s^{-1}$)
PRORP1	Cys-Mito	0.037 ± 0.002	0.062 ± 0.005	670 ± 230	96 ± 10
	Phe-Chlor	0.035 ± 0.002	0.042 ± 0.005	140 ± 50	410 ± 60
	Cys-Nuc	0.037 ± 0.003	0.040 ± 0.002	550 ± 50	73 ± 5
	Phe-Nuc	0.078 ± 0.003	0.035 ± 0.003	160 ± 50	220 ± 60
PRORP2	Cys-Mito	0.013 ± 0.003	0.013 ± 0.002	340 ± 60	37 ± 6
	Phe-Chlor	0.018 ± 0.003	0.015 ± 0.002	340 ± 100	45 ± 10
	Cys-Nuc	0.027 ± 0.002	0.030 ± 0.002	940 ± 130	32 ± 3
	Phe-Nuc	0.035 ± 0.002	0.023 ± 0.002	250 ± 50	95 ± 16
PRORP3	Cys-Mito	0.023 ± 0.002	0.022 ± 0.002	430 ± 30	53 ± 3
	Phe-Chlor	0.023 ± 0.002	0.013 ± 0.002	440 ± 50	32 ± 3
	Cys-Nuc	0.030 ± 0.002	0.008 ± 0.002	420 ±100	17 ± 4
	Phe-Nuc	0.072 ± 0.003	0.062 ± 0.023	2000 ± 850	37 ± 8

[a] Reactions contained 5 μM PRORP, 30 nM pre-tRNA, 30 mM MOPS pH 7.8, 150 mM NaCl, 1 mM TCEP, and 1 mM $MgCl_2$ at 25ºC. The standard error from fitting is reported.
[b] As described in the legend of Fig. 2-4 and Materials and Methods. The standard error from fitting is reported.
[c] Because both correct and incorrect cleavage products are catalyzed with some substrates (see Fig. 2-6, Fig. A-2), k_{max} represents the single-exponential fit to the time course of total product (C_0 + M_{-1}) formation. The multiple turnover kinetics similarly reflect total product formation.

PRORPs bind individual pre-tRNAs similarly

To further explore determinants of substrate recognition, we measured the dissociation constants (K_D) of PRORP1, 2, and 3 for the four *A. thaliana* pre-tRNA substrates using the fluorescence anisotropy binding assay (Fig. 2-5A) (29). The affinity of the PRORPs for these substrates varies by as much as 100-fold (60–6000 nM) (Table 2-3; Fig. 2-5B), demonstrating significant discrimination between the substrates. However, comparison of the dissociation constants reveals that all three PRORPs have comparable binding affinities for a given substrate (less than fourfold difference). All three PRORPs demonstrate the weakest affinity for the Cys-Nuc pre-tRNA and the highest affinity for the Phe-Chlor pre-tRNA. Thus, the PRORP enzymes use similar binding modes for pre-tRNA substrate recognition.

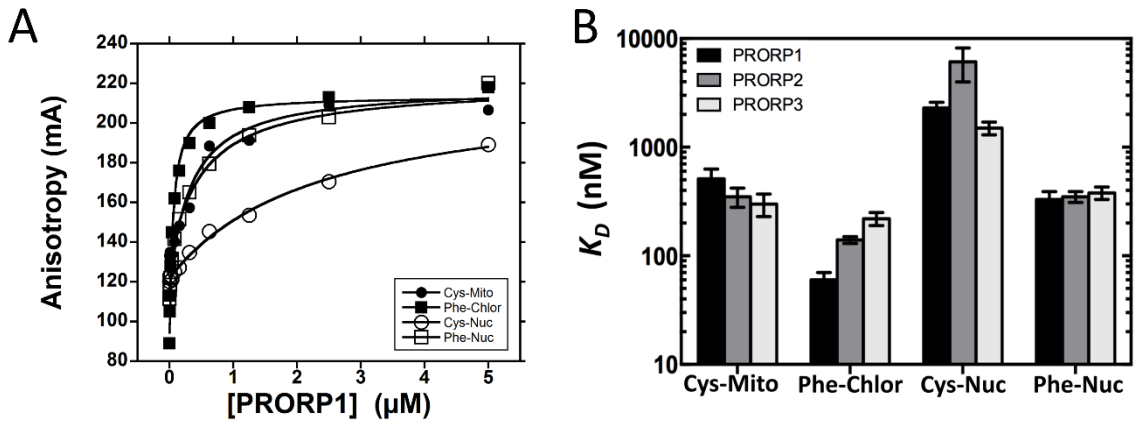


Figure 2-5: (**A**) Representative fluorescence anisotropy binding isotherms for varying concentrations of PRORP1 and 20 nM fluorescein-labeled pre-tRNA (Cys-Mito [closed circle], Phe-Chlor [closed square], Cys-Nuc [open circle] and Phe-Nuc [open square]) in 20 mM MOPS pH 7.8, 300 mM NaCl, 1 mM TCEP, and 1 mM $CaCl_2$. (**B**) Bar graph comparing the binding affinity (K_D) of the PRORP isozymes for four pre-tRNAs. The error bars represent the standard error from fitting.

Table 2-3. Dissociation constants (K_D in nM) for PRORP1, 2, and 3 binding to pre-tRNAs[a]

	pre-tRNA			
Enzyme	Cys-Mito	Phe-Chlor	Cys-Nuc	Phe-Nuc
PRORP1	510 ± 120	60 ± 10	2300 ± 300	330 ± 60
PRORP2	350 ± 70	140 ± 10	6100 ± 2100	350 ± 40
PRORP3	300 ± 70	220 ± 30	1500 ± 200	380 ± 50

[a] Measured as described in the legend of Fig. 2-5 and Materials and Methods. The standard error from fitting is reported.

Product release is not rate-limiting for MTO catalysis by PRORP

The kinetic parameters measured under MTO conditions include all steps in the PRORP kinetic mechanism. To begin dissecting the contribution of discrete steps in the mechanism (e.g., substrate cleavage, product release, etc.) to the progression of the reaction, we measured the STO kinetics of PRORP1, 2, and 3 for catalysis of pre-tRNA cleavage. In contrast to the MTO k_{cat}, the STO rate constant k_{obs} includes only the steps in the kinetic mechanism prior to and including substrate cleavage. For all three PRORPs, the observed STO rate constants (k_{obs}) range between 0.02–0.07 s^{-1} (Table 2-2), regardless of the substrate assayed. The observed rate constant for total product formation (C_0 + M_{-1}) is reported for pre-tRNA substrates that are significantly miscleaved (see following section).

For *B*. *subtilis* RNase P, a significant difference (~ 300-fold) between the STO and MTO rate constants previously revealed that product release is the rate-limiting step in the reaction (30). However, the PRORP MTO k_{cat} values are within twofold of the STO k_{obs} values, suggesting that the rate-limiting step at saturating PRORP is a step prior to product dissociation. The one exception is the difference between the k_{obs} and k_{cat} values for PRORP3 catalyzing cleavage of Cys-Nuc pre-tRNA (k_{obs} is 3.6-fold greater than k_{cat}), suggesting that product release is partially rate limiting for this enzyme/substrate pair.

PRORPs have varying cleavage fidelities

Gel analysis of the STO cleavage assays indicate two distinct product bands in all PRORP-catalyzed reactions with the Phe-Nuc substrate, and for PRORP2 and 3 reactions with the Cys-Mito and Cys-Nuc substrates (Fig. 2-6; Fig. A-2). To further investigate the miscleavage, assay products were separated by high resolution urea-PAGE and compared to *B*. *subtilis* RNase P-catalyzed cleavage of pre-tRNAs (Fig. 2-6). This analysis revealed that the products represent the correct (5-nt, C_0) and incorrect (4-nt, M_{-1}) cleavage products (Fig. 2-6). Additionally, all enzymes cleave *B*. *subtilis* pre-$tRNA^{Asp}$ correctly, indicating that the miscleavage observed with other substrates is not a result of nuclease contamination. *B*. *subtilis* RNase P cleaves all four *A*. *thaliana* pre-tRNAs correctly; only a distinct 5-nt

product band (<5% miscleaved product) is observed. In contrast, the PRORP enzymes have variable cleavage fidelities, with PRORP3 displaying the lowest cleavage fidelity among the PRORPs (Fig. 2-6; Table A-2). Furthermore, the nuclear-localized PRORP2 and PRORP3 significantly miscleave the Cys-Mito and Cys-Nuc substrates whereas the organellar PRORP1 do not (<5%), demonstrating differences in cleavage site selection between the nuclear and organellar PRORPs (Fig. 2-6; Table A-2).

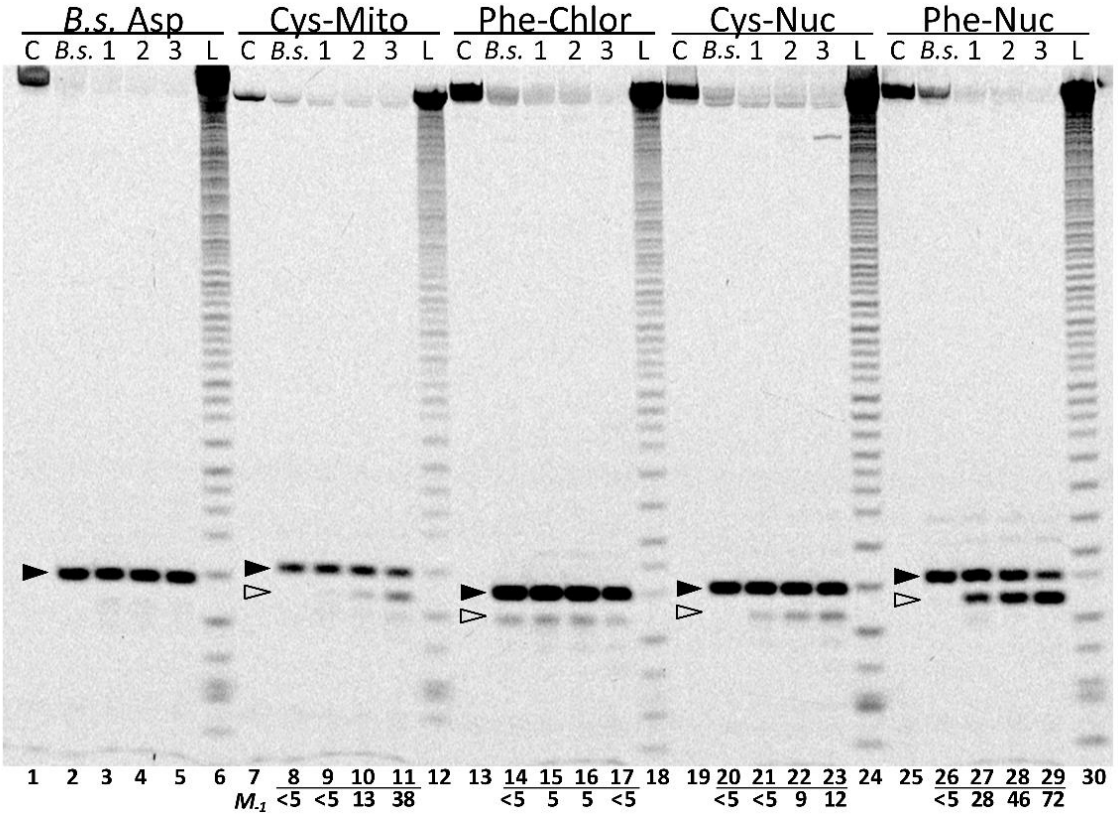


Figure 2-6: PAGE analysis of reaction products from RNase P-catalyzed STO cleavage of pre-tRNA. Of note, 30 nM pre-tRNA substrate was incubated with either 5 µM PRORP for 30 min under standard assay conditions or 1.8 µM *B*. *subtilis* RNase P for 60 min under conditions described in Materials and Methods. The pre-tRNA substrate in each reaction is indicated above the wells. The enzyme in each lane is indicated as follows: (C) no enzyme control, (*B*.s.) *B*. *subtilis* RNase P, (1) PRORP1, (2) PRORP2, (3) PRORP3, and (L) alkaline hydrolysis ladder. The normalized percent miscleaved (% M_{-1}) is indicated below the lane number at the bottom of the gel. Closed arrows indicate the correct cleavage product (5-nt, C_0) and open arrows indicate miscleaved product (4-nt, M_{-1}).

PRORP-catalyzed miscleavage of pre-tRNA between the -2 and -1 nt (M_{-1}) generates a product with a 1-nt leader, a potential substrate for PRORP. Since we are assaying cleavage using a 5′ labeled pre-tRNA, we cannot determine whether the miscleaved product is further processed by PRORP. Thus, we performed

primer extension assays on tRNA at various time-points taken during STO cleavage reactions catalyzed by PRORP1 and 3 of Phe-Nuc pre-tRNA (the substrate miscleaved by all isozymes) (Fig. A-3). These data demonstrate that the miscleaved product is further processed by PRORP to generate mature tRNA (Fig. A-3). However, the observed rate constant for PRORP catalyzed removal of the 1-nt leader is decreased (~ 0.0006–0.001 s^{-1}) (Fig. A-3).

3’ discriminator base can contribute to PRORP cleavage fidelity

One feature of the Phe-Nuc substrate that could lead to the observed miscleavage by all three PRORP enzymes is the formation of a base pair between the adenine discriminator base and the uridine base in the N_{-1} position of the leader (Fig. 2-3), which would extend the acceptor stem helix creating a 4-nt leader. Consistent with this hypothesis, the three pre-tRNA substrates that can form an extra base pair (U/A) between the discriminator base and the N_{-1} leader position (Phe-Nuc, Cys-Mito, and Cys-Nuc) exhibit miscleavage, albeit to varying extents depending on the PRORP (Fig. 2-6). To test this hypothesis, we assessed the ability of PRORPs to correctly cleave three different Phe-Nuc pre-tRNAs variants possessing 5-nt leaders and variable 3’ ends: with (5:1) or without (5:0) a 3’ discriminator base and with an extended 20 nt trailer sequence (5:20) (Fig. 2-7). Removal of the discriminator base (5:0) from Phe-Nuc pre-tRNA eliminated the miscleavage catalyzed by all of the PRORP enzymes, whereas addition of a longer trailer sequence had no significant effect on miscleavage (Fig. 2-7). Similarly, removal of the discriminator base from the Cys-Mito pre-tRNA, to delete a potential A/U base pair with N_{-1}, significantly reduced miscleavage catalyzed by PRORP3 (Fig. A-4).

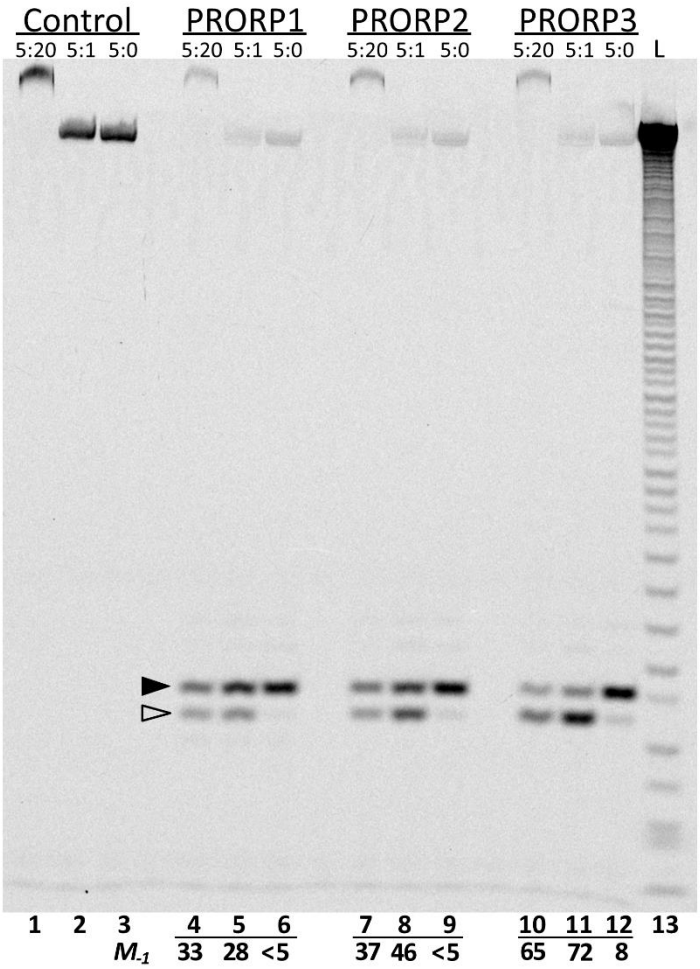


Figure 2-7: Miscleavage of pre-tRNA catalyzed by PRORP is alleviated by removal of the discriminator base. Of note, 30 nM Phe-Nuc pre-tRNA substrates (5:20, 5:1, and 5:0) were incubated with 5 μM PRORP1, 2, or 3 for 30 min under standard reaction conditions. Products were resolved on a 20% gel. Closed arrows indicate the correct cleavage product (5-nt, C_0) and open arrows indicate miscleaved product (4-nt, M_{-1}). The normalized percent miscleaved (% M_{-1}) is indicated below the lane number at the bottom of the gel. In each case, miscleavage is decreased for the substrate lacking the discriminator base (5:0).

Taken together, these data provide evidence that the presence of a base pair between N_{-1} and the discriminator base can engender miscleavage by PRORPs, possibly due to extension of the acceptor stem. However, the potential to form this base pair does not guarantee miscleavage; the Cys-Mito and Phe-Chlor substrates, which contain a potential base pair between the N_{-1} and the discriminator base, are not significantly miscleaved by PRORP1 or 2. Thus, other determinants including sequeneqce context, stability of the pre-tRNA structure, and differences in PRORP substrate recognition must also influence cleavage fidelity.

DISCUSSION

To define the molecular determinants of PRORP recognition, we examined the reactivity of PRORP1 with model substrates and pre-tRNAs with varying leader and trailer lengths. This analysis revealed that PRORP1 makes little or no catalytically important interactions with the trailer or with the leader past the first nucleotide (Table 2-1; Fig. 2-2). This is in contrast to bacterial RNase P where pre-tRNA binding affinity increases ~ 50-fold upon increasing the leader from 1–5 nt (31). Thus, PRORP enzymes do not share a major determinant of molecular recognition utilized by bacterial RNase P. Furthermore, PRORP1 and *B*. *subtilis* RNase P bind mature tRNA ~ 30- and 400-fold weaker than pre-tRNA (leader length ≥ 4-nts), respectively (22). This is consistent with PRORP1 making fewer interactions with the leader than *B*. *subtilis* RNase P. Similar to PRORP1, recent binding studies with *A*. *thaliana* nuclear pre-tRNA^{Gly} suggest PRORP2 does not require long leader (≤ 8-nt) and trailer (1-nt) lengths for tight binding (32). PRORPs efficiently catalyze removal of leader sequences from Cys-Mito tRNA lacking an anti-codon stem-loop (ΔAC) but not from a model substrate lacking the anticodon stem-loop and D-arm (SL substrate). However, bacterial RNase P can catalyze processing of minimal stem-loop model substrates (33), demonstrating a more important role for the interaction between the D-arm and PRORP compared to bacterial RNase P. Thus far, the available data reveal that the molecular determinants of PRORP1 substrate recognition lie within the elbow of the tRNA body and the N_{-1} nucleotide of the leader. Given the similarities in binding affinities and kinetic parameters of the PRORP isozymes for pre-tRNAs, it is likely that PRORP2 and 3 also share these determinants.

The detailed kinetic comparison of *A*. *thaliana* PRORP enzymes reveals both similarities and differences. In general, the three *A*. *thaliana* PRORPs catalyze 5’ end cleavage with comparable catalytic efficiencies (10^4–10^5 M^{-1} s^{-1}) and pre-tRNA binding affinity. This suggests that the differentially localized PRORP isozymes are not selective for pre-tRNAs of different organellar origin. The STO cleavage rate constants among the PRORP isozymes are consistent with a previous study using

a bacterial pre-tRNA^{Gly} substrate, where the observed rate constants vary between 0.04 and 0.13 s^{-1} (34). Despite these similarities, our data indicate that there is a subset of pre-tRNAs that are more efficiently processed by specific PRORP isozymes. For example, PRORP1 has a ~ 10-fold higher k_{cat}/K_M value for cleaving Phe-Chlor pre-tRNA compared to PRORP2 and 3. Nucleotide variations in the D- and TψC-loops may contribute to preferential cleavage and recognition by PRORP1. However, more detailed studies are needed to determine the molecular interactions that confer this enhanced reactivity.

While PRORPs bind and cleave substrates with similar K_D values and cleavage rate constants, they exhibit unexpected differences in cleavage fidelity. For instance, PRORP3 catalyzes the miscleavage of Cys-Mito, Cys-Nuc, and Phe-Nuc pre-tRNA to a greater extent than PRORP1 and PRORP2 (Fig. 2-6). Removal of the discriminator base from Cys-Mito or Phe-Nuc pre-tRNAs results in increased fidelity for all three enzymes, suggesting that the discriminator base engenders miscleavage (Fig. 2-7). Base-pairing between the discriminator nucleotide and the N_{-1} nucleotide of the leader, extending the acceptor stem by one base pair, could account for the observed miscleavage. The greater fidelity observed with PRORP1 could originate from specific interactions with the N_{-1} of the leader and/or interactions with the discriminator base. These data suggest a difference in substrate recognition among the PRORPs with regards to cleavage site selection. While the frequency of PRORP-catalyzed pre-tRNA miscleavage *in vivo* is unknown, PRORP catalyzes phosphodiester bond hydrolysis between -1 nt and +1 nt (correct) and between -2 nt and -1 nt (miscleavage) both *in vivo* and *in vitro* with the atypically processed plant mitochondrial pre-tRNA^{His} (35). Furthermore, PRORPs can correct miscleavage by catalyzing the removal of the miscleaved 1-nt leader, albeit slower than the rate constant for the initial cleavage step (Fig. A-3). This step is likely slow due to base-pairing between the -1 nt and the discriminator base. Cleavage fidelity *in vivo* could be enhanced by RNA binding proteins that decrease incorrect base-pairing or that stabilize RNA structure.

While all three PRORP enzymes catalyze miscleavage of the Phe-Nuc pre-tRNA (Fig. 2-6), *B*. *subtilis* RNase P catalyzes only the correct product from the same substrate. This observation suggests that PRORP enzymes use different criteria to recognize the substrate cleavage site than bacterial RNase P. Bacterial RNase P interacts with the elbow of tRNA (nucleotides within the TψC and D-loops), the acceptor stem, nucleotides within the leader sequence, and the CCA of the 3′ end of pre-tRNA, though all are not required for accurate cleavage (24, 36). Interactions that could increase fidelity of bacterial RNase P compared to PRORP in cleavage of the Phe-Nuc substrate are the base-pairing between the RNA component and both the N_{-1} uridine and the 3′ end of pre-tRNA, leading to splaying of the 3′ and 5′ ends in the bound complex (24, 36). Future experiments examining the nucleotide identity near the cleavage site, including the discriminator base, 1 nt/72 nt base pair, and 3′-CCA, will further define PRORP cleavage site selection.

Based on the kinetic characterization performed here, the substrate selectivity of PRORP2 and 3 are more similar to one another than to PRORP1, consistent with their higher sequence similarity and knockout data that suggest a redundant function for PRORP2 and 3 within the nucleus (12). Furthermore, the fidelity of PRORP2 is more similar to PRORP3 and, interestingly, both enzymes share higher activity toward the SL substrate than PRORP1 (k_{obs} > 12-fold) (Table A-3). This suggests that the nuclear PRORPs have subtle but distinct differences in substrate recognition as compared to the organellar PRORP1. The overlapping substrate specificities of *Arabidopsis* PRORP enzymes suggest that they are in the early stages of diversification, which may be the result of relatively recent gene duplication events (4, 37). It is interesting to note that *Arabidopsis* encodes four variants of the nuclease that catalyzes 3′ end pre-tRNA maturation, tRNase Z (38). These enzymes are differentially localized, but only the chloroplast-localized tRNase Z knockout is lethal, suggesting several isozymes have a redundant function, as with PRORP2 and 3. The abundance of tRNase Z enzymes in *Arabidopsis* is attributed to differential cellular compartmental localization,

differential expression, and/or tissue specific expression (38). The data available suggest that this is also the case with PRORP enzymes.

MATERIALS AND METHODS

Enzyme and substrate preparation

PRORP constructs (Δ76 PRORP1, full-length PRORP2, and Δ9 PRORP3 (4)) were cloned, expressed and purified from *E*. *coli* as previously described (15). The concentrations of PRORP1, 2 and 3 were determined by absorbance using extinction coefficients in the native state at 280 nm of 84,630 M^{-1} cm^{-1}, 91,300 M^{-1} cm^{-1}, and 84,700 M^{-1} cm^{-1}, respectively. Purified enzymes were aliquoted, flash-frozen, and stored at -80°C in 20 mM MOPS pH 7.8, 100 mM NaCl, and 1 mM TCEP. Truncation of PRORP1 was required to obtain soluble protein. However, full-length PRORP2 expressed in *E*. *coli* is observed mainly in the soluble fraction. A Δ20 amino acid truncation of PRORP2 at the N-terminus (the comparable truncation to Δ76 PRORP1) has comparable activity to full-length PRORP2 (32).

Pre-tRNA substrates were prepared by *in vitro* transcription catalyzed by T7 polymerase (39, 40). Substrate sequences were retrieved from the *A*. *thaliana* genomic tRNA database (41). The DNA template for transcription was created by PCR amplification of DNA purchased from Life Technologies GeneArt. To generate fluorescein-labeled pre-tRNA substrates, transcription reactions were performed in the presence of guanosine monophosphorothioate and the 5' phosphorothioate was reacted with 5-(iodoacetamido) fluorescein to label the 5' end, as previously described (27). The pre-tRNA substrates were purified by denaturing PAGE (12%). The discriminator base on the SL substrate was removed to mitigate miscleavage products catalyzed by PRORP2 and 3. The proposed tRNA secondary structures were generated with tRNAscan-SE 1.21 (42).

Multiple-turnover assays

Multiple-turnover (MTO) reactions were performed in a 96-well plate format using a fluorescence anisotropy (FA) assay (as described in Liu *et al*. 2014). Standard reaction conditions (25°C, 30 mM MOPS pH 7.8, 150 mM NaCl, 1 mM

TCEP, and 1 mM $MgCl_2$) were used with an enzyme concentration of 20 nM (PRORP1 and PRORP3) or 80 nM (PRORP2). Higher concentrations of PRORP2 were required to obtain consistent v_0/[E] values at a given substrate concentration. The concentration of fluorescently labeled pre-tRNA was held constant at 40 nM in the reactions while the concentration of unlabeled pre-tRNA substrate was varied. The ratio of labeled to unlabeled substrate did not alter the measured initial rates (data not shown). Black Corning half-area 96-well plates were used with a final reaction volume of 40 µL per well. Initial rates were calculated from the linear decrease in anisotropy (29). The steady-state kinetic parameters were calculated from a fit of the Michaelis-Menten equation (Equation 2-1) to the concentration-dependence of the initial rates using KaleidaGraph (Synergy Software). These kinetic parameters encompass both incorrect and correct cleavage product formation. The kinetic parameters and standard error reported were determined by fitting Equation 2-1 to the concentration-dependence of the initial rates.

$$\frac{v_0}{[\mathrm{E}]} = \frac{k_{\mathrm{cat}}[\mathrm{S}]}{K_{\mathrm{M}}+[\mathrm{S}]} \qquad (2\text{-}1)$$

Anisotropy binding assays

Binding assays were performed as described previously (15). Briefly, the concentration of fluorescein-labeled pre-tRNA was maintained at 20 nM, while the concentration of PRORP was varied (0.005–20 µM). Binding experiments were performed in 20 mM MOPS pH 7.8, 300 mM NaCl, 1 mM TCEP, and 1 mM $CaCl_2$ in a 96-well plate format. PRORP1 can bind but not cleave substrates in $CaCl_2$ (15). Fluorescence anisotropy was measured at an excitation wavelength of 488 nm and an emission wavelength of 535 nm. The observed anisotropies in binding assays where the total fluorescence intensity increased > 15% upon PRORP addition (*B*. *subtilis* substrates containing 1-, 2-, 3-, and 4-nt leaders) were corrected (43). Equation 2-2 was fit to the dependence of the anisotropy on the protein concentration where A is the observed anisotropy, A_0 is the initial anisotropy, ΔA is the total change in anisotropy, [P] is the concentration of PRORP, and K_D is the apparent dissociation constant.

$$A = A_0 + \frac{\Delta A[\mathrm{P}]}{[\mathrm{P}]+K_\mathrm{D}} \quad (2\text{-}2)$$

Single-turnover assays

For single-turnover (STO) reactions, the enzyme and pre-tRNA concentrations were 5 μM and 30 nM, respectively, unless otherwise noted. Reactions were performed in standard conditions (see MTO), initiated by addition of enzyme, and quenched at specified time points (0–1200 sec) with an equal volume of 100 mM EDTA, 6 M urea, 0.1% bromophenol blue, 0.1% xylene cyanol, and 2 μg/μL yeast tRNA. The fluorescently labeled 5' leader product was separated from pre-tRNA by electrophoresis on 20% or 22.5% denaturing PAGE gel. Gels were visualized using a Typhoon 9410 scanner and the fraction product quantified using ImageQuant 5.2 software. The observed single-turnover rate constant was calculated from a fit of a first order exponential equation to the data using KaleidaGraph fitting software (Equation 2-3), where *A* is the end point, *B* is the amplitude, *k* is the observed rate constant, and *t* is time. The STO assays with *Bacillus subtilis* RNase P were performed at 25°C with 1.8 μM bacterial RNase P, 50 mM Tris pH 7.4, 100 mM KCl, 75 mM NaCl, 10 mM DTT, and 3 mM $MgCl_2$. The *B*. *subtilis* RNase P was prepared as previously described (31). The alkaline ladder was produced by the incubation of a given pre-tRNA in 10 mM NaOH and 1 mM EDTA at 95°C for 2 min. Free fluorescein runs as a blur near the second nt (Fig. 2-6). The miscleavage product (4 nt) does not align with the 4-nt ladder product (Fig. 2-6). This may be a result of running near the gel front in combination with differences in the alkaline cleavage products (3' Phosphate) and RNase P products (3' hydroxyl). STO miscleavage kinetics catalyzed by the PRORPs were analyzed as described previously (36, 44). Briefly, the observed rate constants for both correct and miscleaved products are obtained from a single-exponential fit to the data (Equation 2-3). The resulting amplitude (A_c or A_m) for each respective fit is multiplied by the observed rate constant ($k_{obs,c}$ or $k_{obs,m}$) to obtain k_c and k_m (Equations 2-4 and 2-5).

$$Fraction\ product = A - B(e^{-kt}) \quad (2\text{-}3)$$

$$k_c = k_{obs,c}(A_c) \quad (2\text{-}4)$$

$$k_m = k_{obs,m}(A_m) \quad (2\text{-}5)$$

Primer extension

Processing of the Phe-Nuc miscleavage product catalyzed by PRORP1 and PRORP3 was measured by primer extension using Omniscript (QIAGEN) reverse transcriptase (RTase). Substrate was incubated with the PRORPs under standard conditions. Time points were mixed 1:1 with a quench solution containing 10 mM Tris pH 8.0, 50 mM EDTA, and 8 M urea. RNA was repurified by phenol/chloroform extraction and ethanol precipitation. The ladder was generated by alkaline hydrolysis of substrate in 10 mM NaOH and 1 mM EDTA at 95°C for 10 min, then quenched with HCl. The time points, ladder, and substrate were each used to template reverse transcription (RT) reactions. Each template was melted in the presence of an oligo-DNA primer (IDT, 5'-FAMCCCAACTGAGCTATCC-3') at 95°C for 3 min, then cooled to 37°C. The primed-RNA was then mixed 1:1 with a master mix containing Omniscript RT buffer, dNTPs, SUPERase•In (Thermo Fisher), and Omniscript RTase, and incubated at 37°C for 60 min. Glycerol was added for gel loading and products were separated using denaturing PAGE (22.5%) and then fluorescence of the fluorescein label was imaged with a Typhoon 9410 scanner. The substrate and product bands were quantified using ImageQuant 5.2 software. The fraction miscleaved was calculated with [M_{-1}/(C_0 + M_{-1} + Substrate)]. As has been previously reported, some nontemplated nucleotide addition by the RTase was observed (45).

APPENDIX A

This appendix contains supporting tables and figures for Chapter 2.

Tables

Table A-1: Dissociation constant (K_D) and STO observed rate constant (k_{obs}) for PRORP1 with *A. thaliana* mitochondrial pre-tRNA^{Cys} containing varying trailer lengths.

Trailer	K_D (nM)[a]	Fold	k_{obs} (s^{-1})[b]	Fold
0			0.028	0.8
A	65 ± 5	1	0.037	1
ACCA	ND	> 50	0.0083	0.2
AGGU			0.033	0.9
A-23 nt	70 ± 10	1.1	0.033	0.9

[a] Measured under standard binding conditions, except with 100 mM NaCl.
[b] Measured under standard STO assay conditions.

Table A-2: STO miscleavage kinetics catalyzed by PRORP1, 2, and 3.[a]

Phe-Nuc pre-tRNA

Enzyme	$k_{obs,c}$ (s^{-1})[b]	A_c[c]	$k_{obs,m}$ (s^{-1})[d]	A_m[e]	k_c (s^{-1})[f]	k_m (s^{-1})[f]	k_c/k_m
PRORP1	0.067	0.70	0.083	0.30	0.047	0.025	1.9
PRORP2	0.035	0.49	0.035	0.41	0.017	0.014	1.2
PRORP3	0.048	0.25	0.050	0.69	0.012	0.045	0.3

Cys-Nuc pre-tRNA

Enzyme	$k_{obs,c}$ (s^{-1})	A_c	$k_{obs,m}$ (s^{-1})	A_m	k_c (s^{-1})	k_m (s^{-1})	k_c/k_m
PRORP2	0.027	0.77	0.025	0.1	0.021	0.003	7
PRORP3	0.025	0.82	0.020	0.12	0.021	0.002	10.5

Cys-Mito pre-tRNA

Enzyme	$k_{obs,c}$ (s^{-1})	A_c	$k_{obs,m}$ (s^{-1})	A_m	k_c (s^{-1})	k_m (s^{-1})	k_c/k_m
PRORP2	0.014	0.77	0.009	0.12	0.011	0.001	11
PRORP3	0.023	0.58	0.022	0.36	0.013	0.008	1.6

[a] Measured under standard reaction conditions. Error from fitting less than 20%, omitted for clarity.
[b] The observed rate constant for correct product formation (C_o).
[c] Corresponding amplitude of $k_{obs,c}$.
[d] The observed rate constant for miscleavage (M_{-1}).
[e] Corresponding amplitude of $k_{obs,m}$.
[f] Rate constants calculated as described in methods.

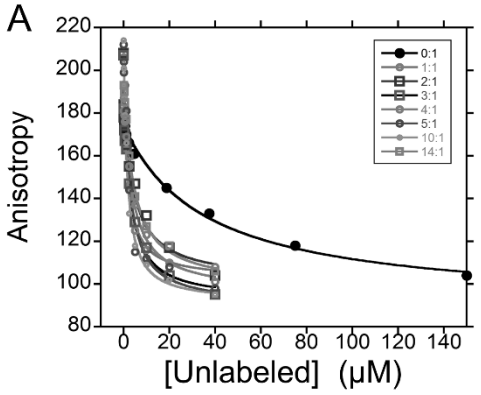
Table A-3: Single-turnover observed rate constants (k_{obs}) for cleavage of model substrates.

	RNA		
Enzyme	Cys-Mito (s^{-1})[a]	Cys-Mito ΔAC (s^{-1})[b]	Cys-Mito SL (s^{-1})[c]
PRORP1	0.037 ± 0.002	0.023 ± 0.005	$< 3.3 \times 10^{-5}$
PRORP2[d]	0.013 ± 0.003	0.012 ± 0.003	$4.2 \pm 0.5 \times 10^{-4}$
PRORP3[d]	0.023 ± 0.002	0.012 ± 0.002	$3.5 \pm 0.2 \times 10^{-4}$

[a] Values from Table 2 shown for comparison.
[b] Measured under standard reaction conditions using gel-based assay.
[c] Measured under standard assay conditions except with 10 μM enzyme using gel-based assay.
[d] k_{obs} calculated from total product (C_0 + M_{-1}) formation. PRORP2 and 3 catalyzed miscleavage of the ΔAC and to a lesser extent the SL substrate because it lacks a discriminator base.

Figures

A



B

Leader	Direct Measurement $K_{D,Fl}$ (nM)[a]	Competition K_D^{app} (μM)[b]	Calculated $K_{D,Un}$ (nM)[c]
0	3400 ± 400	36	5600
1	150 ± 60	2	180
2	310 ± 20	4.3	550
3	140 ± 40	2.8	310
4	150 ± 40	3.5	420
5	190 ± 60	2.3	230
10	100 ± 50	1.7	130
14	100 ± 30	4.2	530

[a] Data from Table 1 shown for comparison.
[b] Obtained by fitting a hyperbola to the data in panel A
[c] Calculated using Equation S1 as in the figure legend.

Figure A-1: Competition binding assays using unlabeled pre-tRNAs. (**A**) ES complex was formed in the presence of 1 mM Ca^{2+} with 5 nM 5:1 *B. subtilis* pre-tRNA and 1 μM PRORP1 (final concentrations). Then unlabeled pre-tRNA was titrated into the reaction and the anisotropy was measured. A hyperbola was fit to the resulting data to obtain an apparent K_D (K_D^{app}) for the unlabeled substrate. (**B**) Affinities for the fluorescein labeled substrate ($K_{D,Fl}$) were measured directly using the anisotropy assay (taken from Table 1). The apparent affinity for unlabeled substrate was measured using the competition assay in panel A (K_D^{app}). The resulting values overestimate the true K_D ($K_{D,Un}$) for the unlabeled substrate. Equation A-1 was used to calculate $K_{D,Un}$ using the values of the $K_{D,Fl}$ of the 5:1 pre-tRNA and the K_D^{app} of each unlabeled pre-tRNA. In Eq. A-1 K_D^{app} is the apparent affinity obtained by the competition assay; $K_{D,Un}$ is the calculated affinity for the unlabeled substrate or tRNA product (*S*); $K_{D,Fl}$ is the affinity for the 5:1 labeled substrate (S_{Fl}); [S_{Fl}] is the concentration of the 5:1 labeled substrate; *f* is the bound/free ratio for S_{Fl} at [S]=0, calculated from equation A-2. In Eq. A-2 f_0 is the fraction of S_{Fl} bound at [*S*]=0, calculated using equation A-3. In Eq. A-3 [*E*] is the concentration of PRORP1.

$$K_{D,Un} = \frac{K_D^{app}}{\left[1+f+\frac{[S_{Fl}]\times(2+f)}{2\times K_{D,Fl}\times(1+f)}\right]} - K_{D,Fl}\times\left(\frac{f}{2+f}\right) \quad \text{(A-1)}$$

$$f = \frac{f_0}{1-f_0} \quad \text{(A-2)}$$

$$f_0 = \frac{([E]+[S_{Fl}]+K_{D,Fl}) - \sqrt{([E]+[S_{Fl}]+K_{D,Fl})^2 - 4\times[E]\times[S_{Fl}]}}{2\times[S_{Fl}]} \quad \text{(A-3)}$$

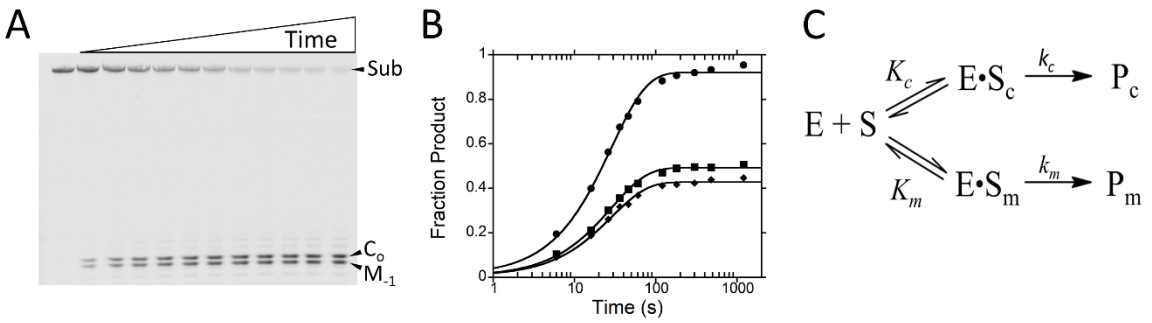


Figure A-2: STO cleavage kinetics of Phe-Nuc pre-tRNA catalyzed by PRORP. (**A**) Representative urea-PAGE of STO cleavage of the Phe-Nuc pre-tRNA catalyzed by saturating PRORP2 under standard reaction conditions. Reactions were quenched at specific time points and resolved on a 22.5% urea-PAGE. C_0 indicates the correct 5-nt product and M_{-1} indicates the miscleaved 4-nt product. (**B**) A plot of the timecourse for product formation. Circles represent the total product formation ((C_0 + M_{-1})/(C_0 + M_{-1} + Substrate)), squares represent correct product formation (C_0 / (C_0 + M_{-1} + Substrate)), and diamonds represent incorrect cleavage (M_{-1} / (C_0 + M_{-1} + Substrate)). A single-exponential equation was fit to the timecourses resulting in values for $k_{obs,t}$, $k_{obs,c}$, and $k_{obs,m}$. (**C**) Scheme for miscleavage, assuming rapid equilibrium binding of pre-tRNA. K_c and K_m represent the equilibrium constants for the respective complex formation, k_c and k_m represent the rate constants for correct cleavage and for miscleavage, respectively. Kinetic parameters for miscleavage are summarized in Table A1.

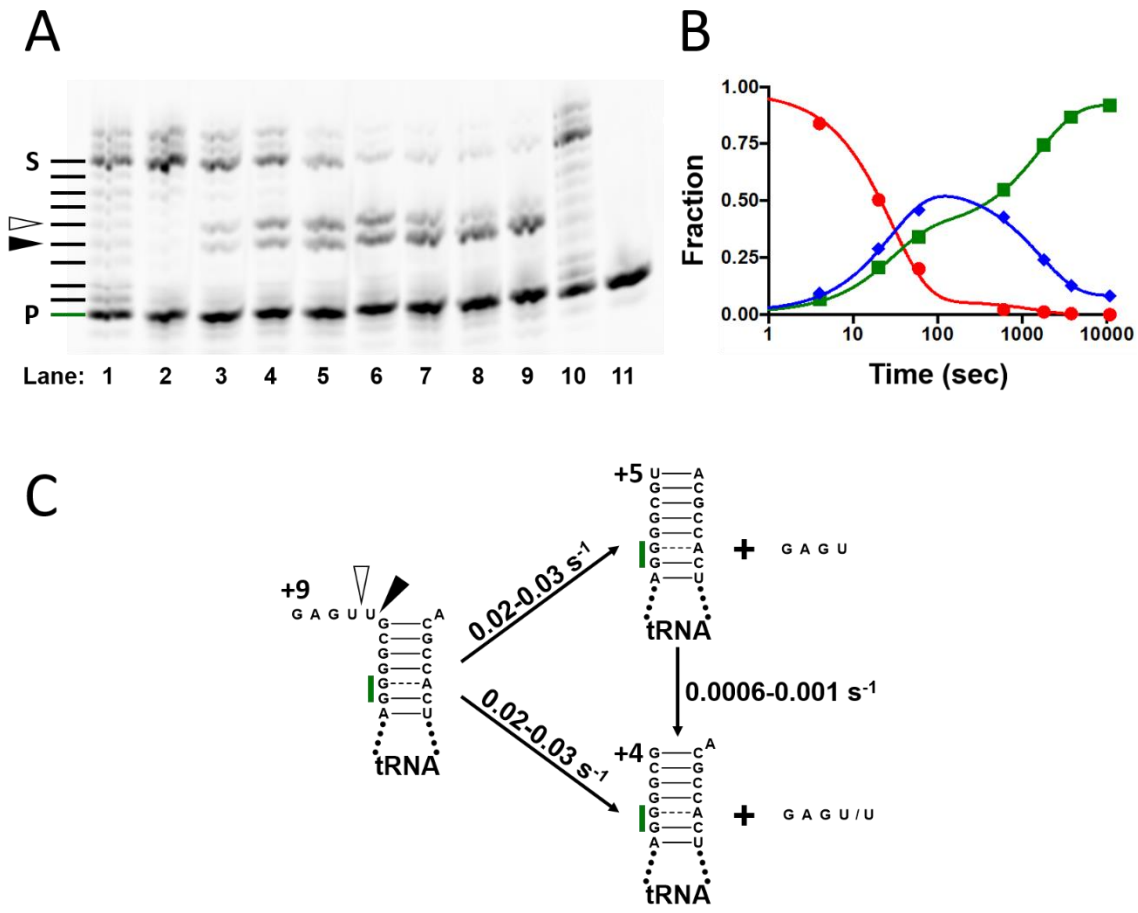


Figure A-3: STO cleavage kinetics of Phe-Nuc pre-tRNA catalyzed by PRORP analyzed by reverse transcription. (**A**) Representative gel of a primer extension assay of the time course of the 5:1 Phe-Nuc pre-tRNA processed by PRORP1 under standard single turnover conditions. Fluorescence of the fluorescein label on the primer is imaged using a Typhoon 9410 scanner. Substrate (S) and primer (P) are indicated, in addition to the correct (+4, closed arrow) and -1 miscleavage (+5, open arrow) products. Lanes are: 1 & 10, ladder; 2, Phe-Nuc pre-tRNA alone; 3–9, cleavage time points of 4 sec, 20 sec, 1 min, 10 min, 30 min, 1 hr, and 3 hr; 11, primer. (**B**) Quantification of the PRORP1 reaction represented as the fraction substrate (circles), product (squares), and miscleavage product (diamonds) at each time point. A double exponential equation was globally fit to the data using GraphPad Prism. Rate constants are $k_{obs,1}$ = 0.03 ± 0.002 s^{-1} and $k_{obs,2}$ = 0.0006 ± 0.00005 s^{-1} (PRORP1) and $k_{obs,1}$ = 0.1 ± 0.008 s^{-1} and $k_{obs,2}$ = 0.001 ± 0.00007 s^{-1} (PRORP3). (**C**) Scheme for processing Phe-Nuc pre-tRNA (left) with miscleavage (top right) and correct (bottom right) products. For clarity, a minimal structure is shown with the acceptor stem and an abbreviated tRNA body. Canonical RNase P cleavage site (closed arrow) and -1 miscleavage site (open arrow) are indicated. Green bars represent the binding site for the 3' end of the oligo-DNA primer. Expected primer extension lengths are given for each molecule. Rate constants k_c and k_m, calculated as described in the methods, are given for the initial cleavage steps.

Figure A-4: Miscleavage of pre-tRNA Cys-Mito catalyzed by PRORP3 is reduced by removal of the discriminator base. Reactions were performed as described in the legend of Fig. 2-6. Closed arrows indicate the correct cleavage product (5-nts, C_0) and open arrows indicate the miscleaved product (4-nts, M_{-1}). The normalized percent miscleaved (% M_{-1}) is indicated below the lane number at the bottom of the gel.

REFERENCES

1. Walker SC, Engelke DR. Ribonuclease P: the evolution of an ancient RNA enzyme. *Crit Rev Biochem Mol Biol*. 2006;41(2):77-102.
2. Howard MJ, Liu X, Lim WH, Klemm BP, Fierke CA, Koutmos M, Engelke DR. RNase P enzymes: divergent scaffolds for a conserved biological reaction. *RNA Biol*. 2013;10(6):909-14.
3. Lechner M, Rossmanith W, Hartmann RK, Thölken C, Gutmann B, Giegé P, Gobert A. Distribution of Ribonucleoprotein and Protein-Only RNase P in Eukarya. *Mol Biol Evol*. 2015;32(12):3186-93.
4. Gobert A, Gutmann B, Taschner A, Gößringer M, Holzmann J, Hartmann RK, Rossmanith W, Giegé P. A single *Arabidopsis* organellar protein has RNase P activity. *Nat Struct Mol Biol*. 2010;17(6):740-4.
5. Holzmann J, Frank P, Löffler E, Bennett KL, Gerner C, Rossmanith W. RNase P without RNA: identification and functional reconstitution of the human mitochondrial tRNA processing enzyme. *Cell*. 2008;135(3):462-74.
6. Vilardo E, Nachbagauer C, Buzet A, Taschner A, Holzmann J, Rossmanith W. A subcomplex of human mitochondrial RNase P is a bifunctional methyltransferase–extensive moonlighting in mitochondrial tRNA biogenesis. *Nucleic Acids Res*. 2012;40(22):11583-93.
7. Vilardo E, Rossmanith W. Molecular insights into HSD10 disease: impact of SDR5C1 mutations on the human mitochondrial RNase P complex. *Nucleic Acids Res*. 2015;43(10):5112-9.
8. Lai LB, Bernal-Bayard P, Mohannath G, Lai SM, Gopalan V, Vioque A. A functional RNase P protein subunit of bacterial origin in some eukaryotes. *Mol Genet Genomics*. 2011;286(5-6):359-69.
9. Taschner A, Weber C, Buzet A, Hartmann RK, Hartig A, Rossmanith W. Nuclear RNase P of *Trypanosoma brucei*: a single protein in place of the multicomponent RNA-protein complex. *Cell Rep*. 2012;2(1):19-25.
10. Hartmann E, Hartmann RK. The enigma of ribonuclease P evolution. *Trends Genet*. 2003;19(10):561-9.
11. Goldfarb KC, Borah S, Cech TR. RNase P branches out from RNP to protein: organelle-triggered diversification? *Genes Dev*. 2012;26(10):1005-9.
12. Gutmann B, Gobert A, Giegé P. PRORP proteins support RNase P activity in both organelles and the nucleus in *Arabidopsis*. *Genes Dev*. 2012;26(10):1022-7.
13. Zhou W, Karcher D, Fischer A, Maximova E, Walther D, Bock R. Multiple RNA processing defects and impaired chloroplast function in plants deficient in the organellar protein-only RNase P enzyme. *PLOS One*. 2015;10(3):e0120533.
14. Sugita C, Komura Y, Tanaka K, Kometani K, Satoh H, Sugita M. Molecular characterization of three PRORP proteins in the moss *Physcomitrella patens*: nuclear PRORP protein is not essential for moss viability. *PLOS One*. 2014;9(10):e108962.
15. Howard MJ, Lim WH, Fierke CA, Koutmos M. Mitochondrial ribonuclease P structure provides insight into the evolution of catalytic strategies for precursor-tRNA 5' processing. *Proc Nat Acad Sci U S A*. 2012;109(40):16149-54.

16. Gobert A, Pinker F, Fuchsbauer O, Gutmann B, Boutin R, Roblin P, Sauter C, Giegé P. Structural insights into protein-only RNase P complexed with tRNA. *Nat Commun*. 2013;4(1353):1-8.
17. Imai T, Nakamura T, Maeda T, Nakayama K, Gao X, Nakashima T, Kakuta Y, Kimura M. Pentatricopeptide repeat motifs in the processing enzyme PRORP1 in Arabidopsis thaliana play a crucial role in recognition of nucleotide bases at TψC loop in precursor tRNAs. *Biochem Biophys Res Commun*. 2014;450(4):1541-6.
18. Howard MJ, Klemm BP, Fierke CA. Mechanistic Studies Reveal Similar Catalytic Strategies for Phosphodiester Bond Hydrolysis by Protein-only and RNA-dependent Ribonuclease P. *J Biol Chem*. 2015;290(21):13454-64.
19. Steitz TA, Steitz JA. A general two-metal-ion mechanism for catalytic RNA. *Proc Natl Acad Sci U S A*. 1993;90(14):6498-502.
20. Emanuelsson O, Brunak S, von Heijne G, Nielsen H. Locating proteins in the cell using TargetP, SignalP and related tools. *Nat Protoc*. 2007;2(4):953-71.
21. Lange A, Mills RE, Lange CJ, Stewart M, Devine SE, Corbett AH. Classical nuclear localization signals: definition, function, and interaction with importin alpha. *J Biol Chem*. 2007;282(8):5101-5.
22. Crary SM, Niranjanakumari S, Fierke CA. The protein component of *Bacillus subtilis* ribonuclease P increases catalytic efficiency by enhancing interactions with the 5' leader sequence of pre-tRNA^{Asp}. *Biochemistry*. 1998;37(26):9409-16.
23. Niranjanakumari S, Stams T, Crary SM, Christianson DW, Fierke CA. Protein component of the ribozyme ribonuclease P alters substrate recognition by directly contacting precursor tRNA. *Proc Natl Acad Sci U S A*. 1998;95(26):15212-7.
24. Reiter NJ, Osterman A, Torres-Larios A, Swinger KK, Pan T, Mondragón A. Structure of a bacterial ribonuclease P holoenzyme in complex with tRNA. *Nature*. 2010;468(7325):784-9.
25. Sun L, Campbell FE, Yandek LE, Harris ME. Binding of C5 protein to P RNA enhances the rate constant for catalysis for P RNA processing of pre-tRNAs lacking a consensus (+ 1)/C(+ 72) pair. *J Mol Biol*. 2010;395(5):1019-37.
26. Hsieh J, Fierke CA. Conformational change in the *Bacillus subtilis* RNase P holoenzyme–pre-tRNA complex enhances substrate affinity and limits cleavage rate. *RNA*. 2009;15(8):1565-77.
27. Rueda D, Hsieh J, Day-Storms JJ, Fierke CA, Walter NG. The 5' leader of precursor tRNA^{Asp} bound to the *Bacillus subtilis* RNase P holoenzyme has an extended conformation. *Biochemistry*. 2005;44(49):16130-9.
28. Bhaskaran H, Taniguchi T, Suzuki T, Suzuki T, Perona JJ. Structural dynamics of a mitochondrial tRNA possessing weak thermodynamic stability. *Biochemistry*. 2014;53(9):1456-65.
29. Liu X, Chen Y, Fierke CA. A real-time fluorescence polarization activity assay to screen for inhibitors of bacterial ribonuclease P. *Nucleic Acids Res*. 2014;42(20):e159.
30. Beebe JA, Fierke CA. A kinetic mechanism for cleavage of precursor tRNA^{Asp} catalyzed by the RNA component of *Bacillus subtilis* ribonuclease P. *Biochemistry*. 1994;33(34):10294-304.

31. Hsieh J, Walker SC, Fierke CA, Engelke DR. Pre-tRNA turnover catalyzed by the yeast nuclear RNase P holoenzyme is limited by product release. *RNA*. 2009;15(2):224-34.
32. Karasik A, Shanmuganathan A, Howard MJ, Fierke CA, Koutmos M. Nuclear Protein-Only Ribonuclease P2 Structure and Biochemical Characterization Provide Insight into the Conserved Properties of tRNA 5' End Processing Enzymes. *J Mol Biol*. 2016;428(1):26-40.
33. Brännvall M, Kikovska E, Wu S, Kirsebom LA. Evidence for induced fit in bacterial RNase P RNA-mediated cleavage. *J Mol Biol*. 2007;372(5):1149-64.
34. Pavlova LV, Gößringer M, Weber C, Buzet A, Rossmanith W, Hartmann RK. tRNA processing by protein-only versus RNA-based RNase P: kinetic analysis reveals mechanistic differences. *ChemBioChem*. 2012;13(15):2270-6.
35. Placido A, Sieber F, Gobert A, Gallerani R, Giegé P, Maréchal-Drouard L. Plant mitochondria use two pathways for the biogenesis of $tRNA^{His}$. *Nucleic Acids Res*. 2010;38(21):7711-7.
36. Zahler NH, Sun L, Christian EL, Harris ME. The pre-tRNA nucleotide base and 2'-hydroxyl at N(–1) contribute to fidelity in tRNA processing by RNase P. *J Mol Biol*. 2005;345(5):969-85.
37. Lynch M, Conery JS. The evolutionary fate and consequences of duplicate genes. *Science*. 2000;290(5494):1151-5.
38. Canino G, Bocian E, Barbezier N, Echeverría M, Forner J, Binder S, Marchfelder A. *Arabidopsis* encodes four tRNase Z enzymes. *Plant Physiol*. 2009;150(3):1494-502.
39. Milligan JF, Uhlenbeck OC. Synthesis of small RNAs using T7 RNA polymerase. *Methods Enzymol*. 1989;180:51-62.
40. Brunelle JL, Green R. *In vitro* transcription from plasmid or PCR-amplified DNA. *Methods Enzymol*. 2013;530:101-14.
41. Michaud M, Cognat V, Duchêne AM, Maréchal-Drouard L. A global picture of tRNA genes in plant genomes. *Plant J*. 2011;66(1):80-93.
42. Schattner P, Brooks AN, Lowe TM. The tRNAscan-SE, snoscan and snoGPS web servers for the detection of tRNAs and snoRNAs. *Nucleic Acids Res*. 2005;33(Web Server issue):W686-9.
43. Lakowicz RJ. *Principles of fluoresence spectroscopy*. Springer, New York. 2006.
44. Loria A, Pan T. Recognition of the 5' leader and the acceptor stem of a pre-tRNA substrate by the ribozyme from *Bacillus subtilis* RNase P. *Biochemistry*. 1998;37(28):10126-33.
45. Chen D, Patton JT. Reverse transcriptase adds nontemplated nucleotides to cDNAs during 5'-RACE and primer extension. *Biotechniques*. 2001;30(3):574-80, 82.

CHAPTER 3

Defining molecular interactions between *Arabidopsis* protein-only Ribonuclease Ps and pre-tRNA[†]

ABSTRACT

Protein-only ribonuclease Ps (PRORP) are enzymes responsible for the 5' end maturation of precursor transfer ribonucleic acids (pre-tRNAs) encoded by various cellular compartments in many eukaryotes. In addition to the PRORP nuclease subunit, metazoan mitochondrial ribonuclease Ps require two additional proteins for efficient catalysis; homologous PRORPs from plants, some protists, and algae act as single-subunit enzymes. Here, we characterize the determinants of substrate binding by the *Arabidopsis thaliana* PRORP1 and PRORP2 using kinetic and thermodynamic experiments. The salt dependence of binding affinity suggests 4–5 contacts with backbone phosphodiester bonds on substrates. These include a single phosphodiester contact with the pre-tRNA 5' leader, consistent with prior reports of short 5' leader requirements. PRORPs contain an N-terminal pentatricopeptide repeat (PPR) domain, truncation of which results in > 30-fold decrease in substrate affinity. Most PPR-containing proteins described to date are implicated in single-stranded RNA binding and can recognize target RNAs in a sequence specific manner. We find that the PPR motifs of PRORPs recognize pre-

[†]This chapter is reformatted from manuscript: Klemm, BP; Karasik, A; Kaitany, KJ; Shanmuganathan, A; Dewar, AJL; Thelen, AZ; Henley, MJ; Jackson, N; Koutmos, M; Fierke, CA. Defining molecular interactions between *Arabidopsis* protein-only Ribonuclease Ps and pre-tRNA. *in preparation*

B.P.K. wrote a draft of the paper. B.P.K. and A.K. designed the mutagenesis experiments. B.P.K., A.K., K.J.K., A.J.L.D., A.Z.T., A.S., N.J., and M.J.H. assayed mutants. A.K. and M.K. performed modeling experiments. B.P.K. and C.A.F. designed the remaining experiments with *At*PRORP1. B.P.K. performed these experiments.

tRNA substrates differently from previously described PPR domains in other proteins. Notably, the PPR domain residues most important for substrate binding in PRORPs do not correspond to positions involved in single-stranded sequence specific base recognition in other PPR proteins. Several of these residues are highly conserved in PRORPs from algae, plants, and metazoans, suggesting a conserved strategy for substrate recognition by the PRORP PPR domain. This work defines several molecular determinants of PRORP-substrate recognition and provides a new predictive model for the PRORP-substrate complex.

INTRODUCTION

Ribonuclease P (RNase P) enzymes are essential endonucleases with diverse macromolecular composition that are responsible for the maturation of the 5’ end of precursor transfer ribonucleic acid (pre-tRNA) (1). In many biological settings, RNase P is a ribozyme with a large catalytic RNA capable of processing pre-tRNAs *in vitro* (2). The ribozyme is associated with one or more protein components required for function *in vivo* (3). RNase P proteins increase substrate affinity and the ability of divalent metal ions to bind at specific sites (4, 5).

In many eukaryotic species, including protists, algae, land plants, and metazoans, protein-only RNase Ps (PRORPs) exist (6-9). Human mitochondrial RNase P (mtRNase P) was the first PRORP described and it requires 2 additional protein subunits for activity (6). These subunits are an m^1G/A_9 tRNA-methyltransferase (TRMT10C, also MRPP1) and a hydroxysteroid dehydrogenase/reductase (HSD17B10, also MRPP2), which form a sub-complex (6, 10). The methyltransferase activity is not required for the RNase P activity (10). Thus, it is proposed that the subunits contribute primarily to substrate recognition.

In contrast to the metazoan PRORP, the PRORPs from algae, protists, and plants do not require additional subunits for efficient catalysis *in vitro* (7-9, 11), suggesting differences in substrate recognition. The three PRORPs from *A. thaliana* are designated PRORP1–3. *At*PRORP1 localizes to the mitochondria and chloroplasts where it is fully responsible for pre-tRNA maturation (7), while

*At*PRORP2 and *At*PRORP3 co-localize to the nucleus and are not fully redundant in nuclear pre-tRNA processing (11). *At*PRORP1 utilizes a 2-metal mechanism similar to the ribozyme, relying on the ionization of metal-bound waters for nucleophile activation in catalysis (12). Given the additional mechanistic information and the relative simplicity of the *At*PRORPs, we have used them as a model system to study PRORP-substrate molecular recognition.

PRORPs contain a unique domain architecture (Fig. 3-1A). An N-terminal pentatricopeptide repeat (PPR) domain significantly enhances the affinity for substrate and truncation of the first 3–4 repeats abolishes catalytic activity (13, 14). Thus, the PPR domain is proposed to both bind and orient substrate with respect to the metallonuclease domain (13). In addition, the nuclease domain is a member of the Nedd4-BP1, YacP nuclease (NYN) family (15). Lastly, a bipartite CC/HC Zn^{2+}-binding domain flanks the NYN domain (13). Our current understanding of how each domain, in particular the PPR domain (Fig. 3-1B), contribute to PRORP substrate recognition is limited.

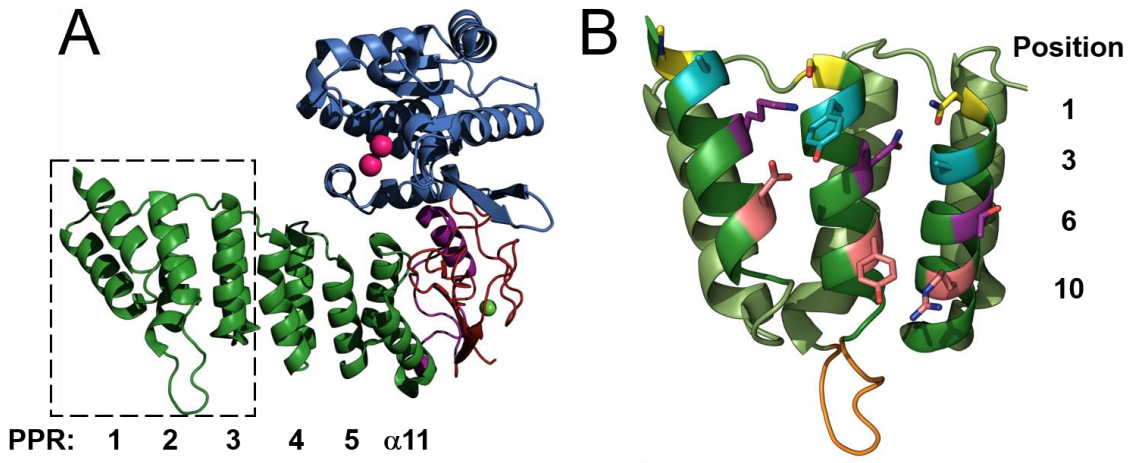


Figure 3-1: Structure of Arabidopsis thaliana PRORP1 (PDB 4g24), generated in PyMol (16). (**A**) Overall architecture of AtPRORP1. PPR motifs are in green and numbered from the N-terminal end, a plant-specific helical insertion in purple, central domain in red, and NYN domain in blue. Mn^{2+} ions displayed as pink spheres and Zn^{2+} ions displayed as a green sphere. The region marked by the dashed box is expanded in panel B. (**B**) AtPRORP1 PPR domain motifs 1–3, generated in PyMol. For each PPR motif, position 1 is colored yellow, position 3 is colored cyan, position 6 is colored purple, and position 10 is colored pink. A loop between PPR motif 2 A and B helices, to which additional Physcomitrella patens PRORP PPR sequence aligns, is colored orange (Fig. B-1).

Several previous results suggest differences in substrate recognition between the bacterial ribozyme and PRORP NYN domain. Unlike the ribozyme, PRORP active site metals apparently do not contact the pro-R_P oxygen of the scissile phosphodiester bond (17). Furthermore, while the 3'-CCA is specifically recognized by the bacterial ribozyme, it is either inhibitory or immaterial to *At*PRORP activity (18, 19). Additionally, *At*PRORP1 and *At*PRORP3 do not significantly contact either the 5' leader sequence beyond N_{-2} or the 3' trailer (Fig. 3-2A); these regions do not alter substrate affinity or catalytic activity (19, 20). The minimal 5' and 3' end interactions indicate that PRORP substrate recognition lies primarily within the tRNA body.

Previously, a nuclease footprinting assay demonstrated that there was significant protection of bases in the D- and TψC-loops (Fig. 3-2A) by *At*PRORP1 (18). Given these data and the likelihood that the NYN domain binds at the scissile phosphodiester bond, it was proposed that the PPRs recognized the pre-tRNA elbow (the structure formed by interaction between the D- and TψC-loops) (18). However, this remains to be confirmed experimentally. Recent attempts to alter base specificity of the PPR domain were unsuccessful (19), although there may be base-specificity that has not been detected. Furthermore, while the TψC-arm is sufficient for recognition and catalysis by plant PRORPs, the presence of a D-arm increases the affinity significantly (19, 20). These data provide a basis for us to examine the substrate features that contribute to recognition.

A model of substrate-bound *At*PRORP1 was previously generated using molecular dynamics, including the PPR domain docked to the TψC-loop (14). The authors assumed that PRORPs use the recognition strategy employed by several single-stranded RNA binding PPR proteins. The ssRNA-binding PPR proteins recognize nucleobases utilizing residues in two tandem repeats at positions 6 and 1', as well as hydrophobic amino acids at position 3 that contribute by van der Waals or stacking interactions (21, 22). Cleavage assays with *At*PRORP1indicated that mutations to position 6 of PPR motifs 2, 3, and 4 reduced activity ≤ 70% (14).

Nonetheless, the full suite of PPR residues that are important for PRORP substrate binding remain to be identified.

Here, we characterize the mode of substrate binding and recognition by the highly conserved *At*PRORP1 and 2 using a variety of biochemical techniques. The salt-dependence of pre-tRNA affinity indicate that *At*PRORP1 and *At*PRORP2 make at least four direct contacts to substrate backbone phosphodiester bonds, including a single phosphodiester contact in the leader. Importantly, these interactions are not sequence specific. The salt-dependence of affinity for mature tRNA indicate that a significant portion of the affinity for substrate stems from interactions to the sugars and/or bases in the body of the substrate, in a contrast to the bacterial ribozyme. To test whether *At*PRORP1 uses canonical PPR-nucleobase interactions, we mutated both residues in the PRORP PPR domains and nucleotides in a pre-tRNA substrate and assessed how the mutations impact the ability of PRORP to bind. In contrast to other known PPR proteins, PRORP does not exhibit demonstrable sequence selectivity for binding its substrates. These experiments provide a biochemical framework for understanding the molecular recognition of complex RNA structures by the non-canonical PPRs of plant PRORPs.

RESULTS

*At*PRORP-substrate recognition mode

To begin characterizing how *At*PRORPs recognize their cognate substrates, we set out to determine the general mode of substrate binding. We first measured the dependence of the protein-nucleic acid interaction on the concentration and identity of ions in solution. These data parse the dependence of affinity on ionic interactions with backbone phosphodiester bonds, compared to that of non-ionic interactions. Monovalent and divalent cations directly interact with backbone phosphodiester bonds on nucleic acids. These ions must be released for a protein to directly contact those sites, thus affinity depends on the cation concentration (23, 24). Cations may also associate with nucleic acids through a thermodynamic

cloud containing both cations and anions, which inhibits protein-nucleic acid interactions through a separate non-specific screening mechanism.

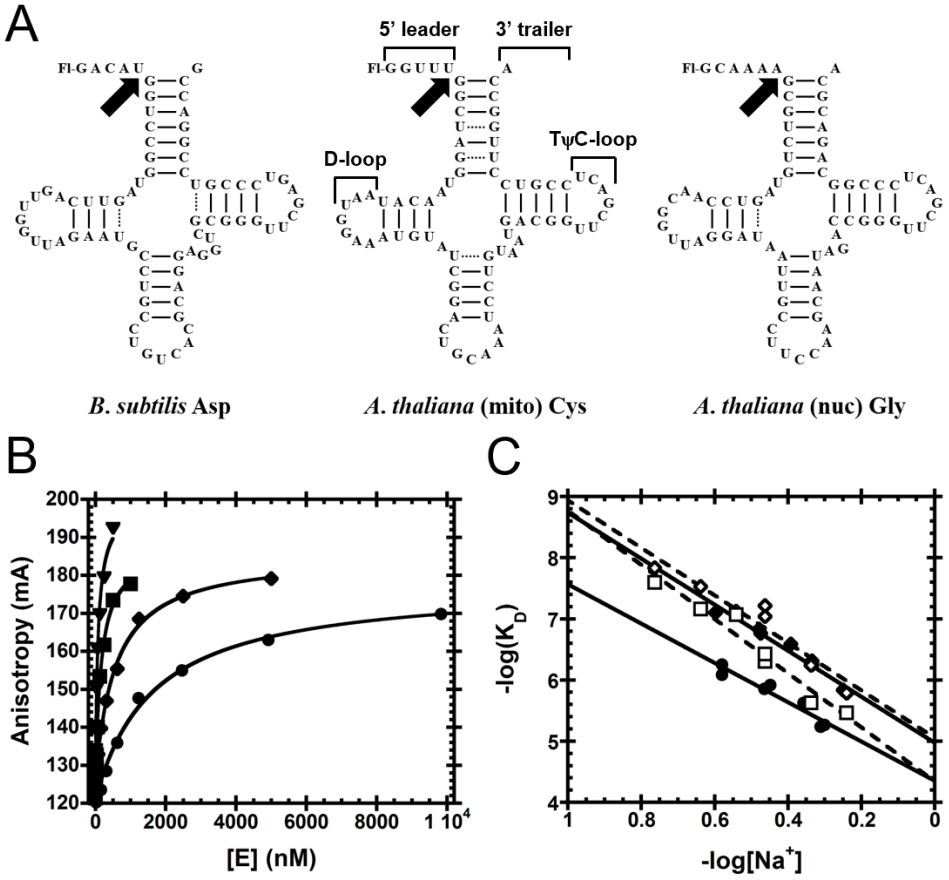


Figure 3-2: Substrates used and thermodynamic assays. (**A**) Substrates used in this manuscript. Pre-tRNA^{Asp} is from *Bacillus subtilis*, while pre-tRNA^{Cys} and pre-tRNA^{Gly} are from the *Arabidopsis thaliana* mitochondrial and nuclear genomes, respectively. Structural features of pre-tRNA are detailed on pre-tRNA^{Cys}, including the 5' leader, 3' trailer, and the D- and TψC-loops. Black arrows indicate canonical RNase P cleavage site. (**B**) Fluorescence anisotropy binding curves for AtPRORP1 binding to *B. subtilis* pre-tRNA^{Asp}. A hyperbola (Equation 3-1, Materials and Methods) was fit to the data. Data were measured in 30 mM MOPS pH 7.8, 1 mM TCEP, and 20 mM $CaCl_2$ with 250 mM (▼), 330 mM (■), 450 mM (♦), and 550 mM NaCl (●). (**C**) Na^+-dependence of AtPRORP binding, plotted as the $-\log(K_D)$ versus $-\log[Na^+]$. Equation 3-4 (Materials and Methods) was fit to the data (25). The slope of the line (Z) reports on the number of ionic interactions made to substrate phosphodiester bonds, while the intercept [$\log(K_0)$] reports on the non-ionic contributions to binding. Data include AtPRORP1 binding to pre-tRNA^{Asp} (♦) and pre-tRNA^{Cys} (●) in 20 mM Ca^{2+}, as well as AtPRORP2 in 6 mM Ca^{2+} binding to pre-tRNA^{Asp} (◊) and pre-tRNA^{Gly} (□).

We established the dependence of *At*PRORP1- and *At*PRORP2-substrate binding affinity on ions in solution to estimate the number of backbone phosphodiester contacts. We determined dissociation constants (K_D) for *At*PRORP1 and 2 by fluorescence anisotropy (FA) assays using a *B. subtilis* pre-tRNA^{Asp} (*At*PRORP1/2), an *A. thaliana* mitochondrial pre-tRNA^{Cys} (*At*PRORP1),

each with 5-nt leaders, and an *A*. *thaliana* nuclear pre-tRNA^{Gly} (*At*PRORP2) with a 6-nt leader; all substrates have a fluorescein label at the 5' end (Fig. 3-2A). The pre-tRNA^{Asp} has been used extensively with the bacterial ribozyme, allowing us to make direct comparisons to PRORPs, while the pre-tRNA^{Cys} is a cognate substrate for *At*PRORP1 and the pre-tRNA^{Gly} is a cognate substrate for *At*PRORP2. We obtained thermodynamic affinities (K_D) by fitting a hyperbola (Equation 3-1, Materials and Methods) to the data (Fig. 3-2B).

The Na^+-dependence of the K_D shows a linear dependence in a log-log plot (Fig. 3-2C), as described by Equation B-1 (Supporting Methods), which was adapted from Equation 18 of deHaseth, *et al*. 1977 (25). Divalent cations are required to fold the pre-tRNA, so $CaCl_2$, which does not activate *At*PRORP1 or *At*PRORP2 (13, 26), was supplied at a constant value for each measurement; 6 or 20 mM $CaCl_2$ for *At*PRORP2 or *At*PRORP1, respectively. We use 6 mM $CaCl_2$ to measure binding with wild type and mutant *At*PRORP2 because the K_D values are above the measurable range in 20 mM $CaCl_2$. The 6 mM $CaCl_2$ conditions resulted in *At*PRORP1 substrate affinities that were at most 65% weaker than the 20 mM conditions, yet altered the slope of the Na^+-dependence < 10% (data not shown). Given that we observed much larger changes in affinity resulting from Na^+-dependence (up to 135-fold between *At*PRORP2 and pre-tRNA^{Gly}), we have continued with the analysis as reported. We observe minimal competition between the Ca^{2+} and Na^+ for the RNA substrate under the concentrations used for the binding assays for *At*PRORP1, as evidenced by the relatively linear Na^+-dependence in the log-log plot (Fig. 3-2C). We maintained constant pH during the experiments and anion effects are precluded based on the $CaCl_2$-alone and Na_2SO_4 data described below. In the absence of these effects, Equation B-1 can be reduced to Equation 3-4 (Materials and Methods), which was fit to the data.

The slope of a -log(K_D) versus -log[Na^+] plot is given by $Z\varphi$, where φ is the fraction of Na^+ associated thermodynamically with each backbone phosphodiester

and Z is the number of phosphodiester bonds the protein contacts.[⊥] Previous data suggested that the φ^{Na} for structured RNAs such as pre-tRNA is comparable to the φ^{Na} for dsDNA (27). Thus, we used dsDNA φ^{Na} = 0.88 in fitting Equation 3-4 to the data, yielding the values in Table 3-1. The Z values for *At*PRORP1 suggest the formation of four protein-phosphodiester contacts upon binding pre-tRNA. The values for *At*PRORP2 are higher, suggesting contacts with five phosphodiester groups.

Table 3-1. Na^+-dependence of binding affinity.

Enzyme	Substrate	Leader	K_D (nM)[a]	Z[b]	log(K_0)[b]	ΔG_0 (kcal/mol)[c]
*At*PRORP1	Asp	5-nt	160 ± 20	4.3 ± 0.3	5.0 ± 0.1	-6.9 ± 0.1
		1-nt	700 ± 80	3.7 ± 0.2	4.66 ± 0.06	-6.4 ± 0.1
		0-nt	19100 ± 600	2.8 ± 0.5	3.5 ± 0.2	-4.8 ± 0.3
	Cys	5-nt	1290 ± 130	3.7 ± 0.5	4.4 ± 0.2	-6.0 ± 0.3
*At*PRORP2	Asp	5-nt	80 ± 20	4.4 ± 0.6	5.0 ± 0.3	-6.9 ± 0.4
	Gly	6-nt	440 ± 90	5.0 ± 0.5	4.3 ± 0.2	-5.9 ± 0.3

[a] Mean and standard deviation reported are from two independent experiments in 330 mM NaCl.
[b] Value and error from fitting Equation 3-4 to the Figure 3-2C data using φ^{Na}= 0.88 as described in the Materials and Methods (25).
[c] Calculated using $\Delta G_0 = -RT \times \ln K_0$.

The estimated affinity at 1 M NaCl has been used to estimate the contribution of non-ionic interactions to the affinity in model systems using normal Gibbs free energy definitions (28). For *At*PRORP1 at 27°C, the log(K_D) at 1 M NaCl indicates values of -6.9 ± 0.1 and -6.0 ± 0.3 kcal/mol for pre-tRNA^{Asp} and pre-tRNA^{Cys}, respectively (Fig. 3-2C). For *At*PRORP2, the values are -6.9 ± 0.2 and -5.9 ± 0.4 kcal/mol for pre-tRNA^{Asp} and pre-tRNA^{Gly}, respectively. Depending on the context, hydrogen bonds can supply 1–3 kcal/mol of free energy, while van der Waals interactions such as base stacking can supply 0.5–2 kcal/mol (29). Thus, *At*PRORPs could make as many as 6 hydrogen bonds, up to 14 van der Waals interactions, or more likely an intermediate combination of both types of interactions with pre-tRNA^{Asp}. The reduced affinity for pre-tRNA^{Cys} or pre-tRNA^{Gly} compared to pre-tRNA^{Asp} is attributed to a loss of approximately 1 kcal/mol of non-ionic interactions.

[⊥] M^+ that are released from the nucleic acid upon binding to the protein, which approximates the number of protein-phosphodiester contacts.

*At*PRORP1 does not have specific anion binding sites that compete with substrate binding

Specific anion binding sites on proteins can decrease binding affinity of nucleic acids, in an ion-dependent manner similar to the effect of cations. Previously, anion sites on proteins have been probed by comparing the dependence of binding affinity on the concentrations of monovalent (M^+) and divalent (M^{2+}) cations for a given anion (24, 25). For a protein binding to dsDNA in the absence of screening by anions, the theoretical ϕ^{Mg}/φ^{Na} is 0.53, which corresponds to the difference in the cations' occupancy on the phosphodiester bonds in the backbone (25). To test whether anion binding contributes to the salt dependence of binding affinity, we measured dissociation constants in the presence of varying concentrations of $CaCl_2$ (alone) or Na_2SO_4.

The pre-tRNA^{Asp} anisotropy in the absence of PRORP and Na^+ displays a hyperbolic dependence on $CaCl_2$ ($K_{D,app}$ = 11 ± 3 mM, Fig. B-2A), with < 10% change in total fluorescence. This is likely due to stabilization of the tRNA structure upon addition of divalent cations (30), as the 5'-fluorescein likely has less rotational freedom when the tRNA structure is compact (31). Above 50 mM Ca^{2+} the anisotropy of free substrate changed less than 20%, so we varied [$CaCl_2$] from 50 to 125 mM and observed a decrease in *At*PRORP1 affinity (Fig. 3-3). Fitting Equation 3-4 to the data with Z = 4, we obtained ϕ^{Ca} = 0.51 ± 0.04. This value is in relatively good agreement with ϕ^{Mg} = 0.47 for dsDNA (25). The ratio of the slopes in Ca^{2+} and Na^+ is 0.54, so the salt-dependence of PRORP binding affinity can be explained using only the thermodynamic occupancies of the cations on backbone phosphodiester bonds. The slope of log(K_D) vs log[Na_2SO_4] is < 20% smaller than the slope of the NaCl data, resulting in tighter binding in Na_2SO_4 at higher Na^+ concentrations (Fig. 3-3). When we re-plot the data against total ionic strength, the NaCl and Na_2SO_4 data overlay, while the $CaCl_2$ data does not overlay with either NaCl or Na_2SO_4 (Fig. B-2B). Thus, the modest differences in affinity between the NaCl and Na_2SO_4 results are most likely due to differences in ionic strength. Given these data, we exclude the term for anion effects from our fits. The 1 M NaCl, 1 M

Na_2SO_4, and 1 M $CaCl_2$ ionic strength intercepts are within error (Fig. B-2B), confirming that the non-ionic contributions to binding are ion-independent.

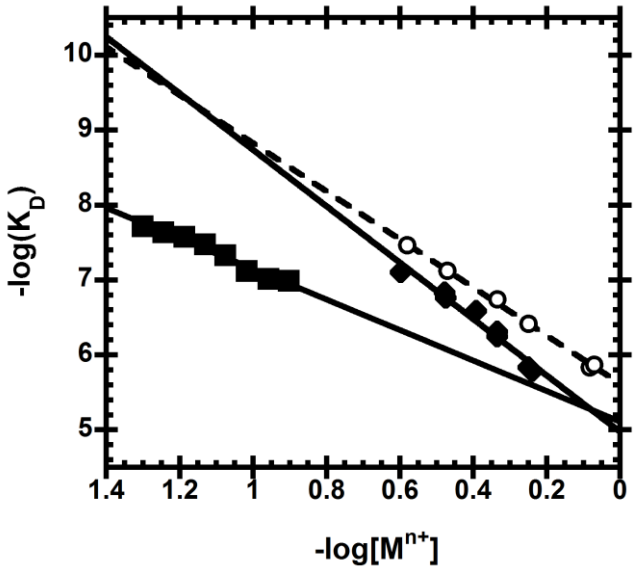


Figure 3-3: Dependence of binding affinity on cations/anions. Cation (M^{n+}) dependence of dissociation constants for AtPRORP1 binding to pre-$tRNA^{Asp}$. Equation 3-4 (Materials and Methods) was fit to the data with the dsDNA φ^{Na} = 0.88 or ϕ^{Mg} = 0.47 (25). Data include AtPRORP1 binding in NaCl (♦, Z = 4.3 ± 0.3, $\log(K_0)$ = 5.0 ± 0.1), Na_2SO_4 (○, Z = 3.6 ± 0.1, $\log(K_0)$ = 5.62 ± 0.03), and $CaCl_2$ (■, Z = 4.3 ± 0.3, $\log(K_0)$ = 5.1 ± 0.2). The slope of the line (Zφ or Zϕ) reports on the number of ionic interactions made to substrate phosphodiester bonds, while the intercept [$\log(K_0)$] reports on the non-ionic contributions to binding.

Cation and anion identities have limited effects on *At*PRORP1 affinity

Given that protein-nucleic acid interactions can be strongly dependent on the identity of the ions in solution, we examined the binding affinity with the 5-nt pre-$tRNA^{Asp}$ for *At*PRORP1 with several different cations. We supplied M^+ as MCl at 330 mM and observe at most a 2.6-fold effect on affinity in the order of $Li^+ < K^+ \approx Rb^+ \approx NH_4^+ < Na^+$ (Table B-1). With the exception of Na^+, this correlates with the order of the cations' affinities for nucleic acids: $Na^+ < Li^+ < K^+ < Rb^+ < NH_4^+$ (32). We next varied the anion (X^{n-}) identity as Na_nX^{n-} at 330 mM Na^+ and observe an 80-fold effect on affinity in the order of $CH_3CO_2^- < SO_4^{2-} < Cl^- < NO_3^- \approx Br^- < I^-$ (Table B-1). The trend is similar to that observed for LacI, which was explained in part by the ions' positions in the lyotropic series, with anions that unfold protein to a greater extent (I^-, Br^-, etc.) resulting in lower affinity (24). However, the effects we observe for *At*PRORP1 are much smaller. The binding defect observed for LacI in I^- was more than 10^4-fold when compared to $CH_3CO_2^-$ (24). Alternatively, it is

possible that the effect we observe is due to pre-tRNA misfolding, given that anion identity can affect RNA structure (33). As stated above, the dependence of binding on $NaSO_4$ was comparable to that of NaCl. Given that we observe small effects of ion identity on *At*PRORP1 binding, the binding determinant trends we define in NaCl are likely to translate well to other buffering/ionic conditions.

*At*PRORP1 makes fewer contacts to substrate leader than the bacterial ribozyme

It was previously shown that varying the leader length of pre-tRNA substrates beyond 1- or 2-nt has little effect on the single-turnover activity and binding affinity with *At*PRORPs (19, 20). From these data, it is apparent that *At*PRORPs can process a substrate with short 1- and 2-nt leaders, and that *At*PRORP1 discriminates against binding the tRNA product (> 30-fold lower affinity for tRNA than pre-tRNA). In contrast, the *B*. *subtilis* RNA-based RNase P relies on extensive contacts with the leader and trailer sequences for substrate recognition and displays a significant dependence on leader length beyond 2 nt (34, 35).

We determined the Na^+-dependence of binding for the fluorescein-labeled 1-nt pre-$tRNA^{Asp}$ and $tRNA^{Asp}$ product to evaluate the nature of the *At*PRORP1 interactions with the leader (Table 3-1, Fig. 3-4). For the 1-nt substrate, the main effect is a value for K_0 that is decreased 2-fold compared to the 5-nt substrate, suggesting a 0.5 kcal/mol reduction in non-ionic interactions with the shorter leader. The value of Z is also reduced modestly (≈ 20%). The intercept of the 1-nt substrate represents a non-ionic contribution to binding of -4.9 ± 0.3 kcal/mol. In contrast, the Z valued for the tRNA product is reduced to 2.8 ± 0.5, consistent with the loss of one phosphodiester bond contact. The value of K_0 also decreases significantly compared with pre-$tRNA^{Asp}$ containing either a 5 nt or 1 nt leader, equivalent to a loss of 2.1 and 1.6 kcal/mol, respectively, indicative on non-ionic interactions with the leader. However, the K_0 for tRNA indicates that non-ionic interactions with tRNA are important determinants of binding affinity, contributing ~5 kcal/mol – over 70% of the total binding energy.

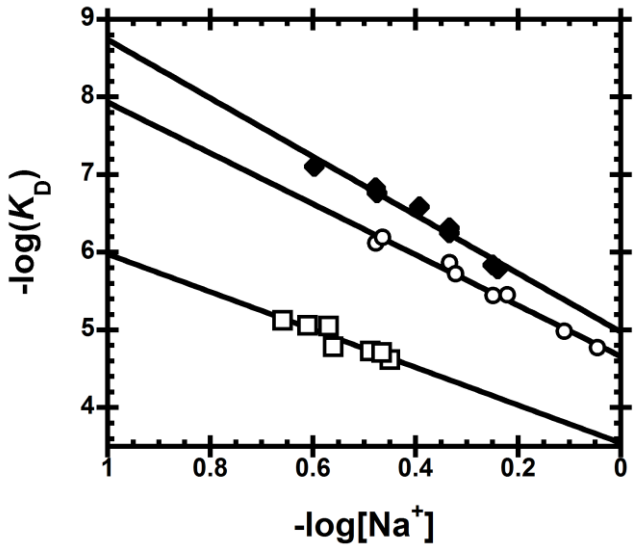


Figure 3-4: Dependence of AtPRORP1 binding affinity on leader length. Na^+-dependence of AtPRORP binding. Equation 3-4 (Materials and Methods) was fit to the data (25). The slope of the line (Z) reports on the number of ionic interactions made to substrate phosphodiester bonds, while the intercept [log(K_0)] reports on the non-ionic contributions to binding. Data include AtPRORP1 binding to 5-nt pre-$tRNA^{Asp}$ (♦, Z = 4.3 ± 0.3, log(K_0) = 5.0 ± 0.1), 1-nt pre-$tRNA^{Asp}$ (○, Z = 3.7 ± 0.2, log(K_0) = 4.7 ± 0.1), and $tRNA^{Asp}$ (□, Z = 2.8 ± 0.5, log(K_0) = 3.5 ± 0.2).

The PRORP PPR domain recognizes tRNAs using non-canonical positions

Previous work demonstrated that mutations of N136T, T180N, and G215N, each at position 6 of an *At*PRORP1 PPR motif, resulted in minor pre-tRNA processing defects (14). To further characterize substrate recognition by the PRORP PPR domain we first screened the pre-tRNA binding affinity and salt-dependence for variants of seven residues in *At*PRORP1 that are 1) highly- or fully-conserved among plant PRORPs, 2) located on the PPR surface facing the NYN domain, and 3) have the potential to make hydrogen-bonding, ionic, or base-stacking interactions (Fig. 3-5). In addition to the residues at 1 and 6 positions that have been shown to be involved in base selection in other PPR domains (21, 22), we also targeted residues at position 10, which were not identified in the canonical base-selection motifs (36). While residues at position 3 in ssRNA binding PPR proteins are typically hydrophobic (i.e., Leu, Phe), in PRORPs the residues at this position are mostly small or hydrophilic. Figure 3-5A shows the position of the residues that we targeted: Tyr 133 (position 3; PPR motif 2), Asn 136 (position 6; PPR2), Tyr 140 (position 10; PPR2), Asn 175 (position 1; PPR3), Thr 180 (position 6; PPR3), Arg 184 (position 10; PPR3), and Arg 212 (position 3; PPR4). We

examined the effects of alanine mutations to each position, as well as more conservative mutations such as Y133F, T180S, and R184K.

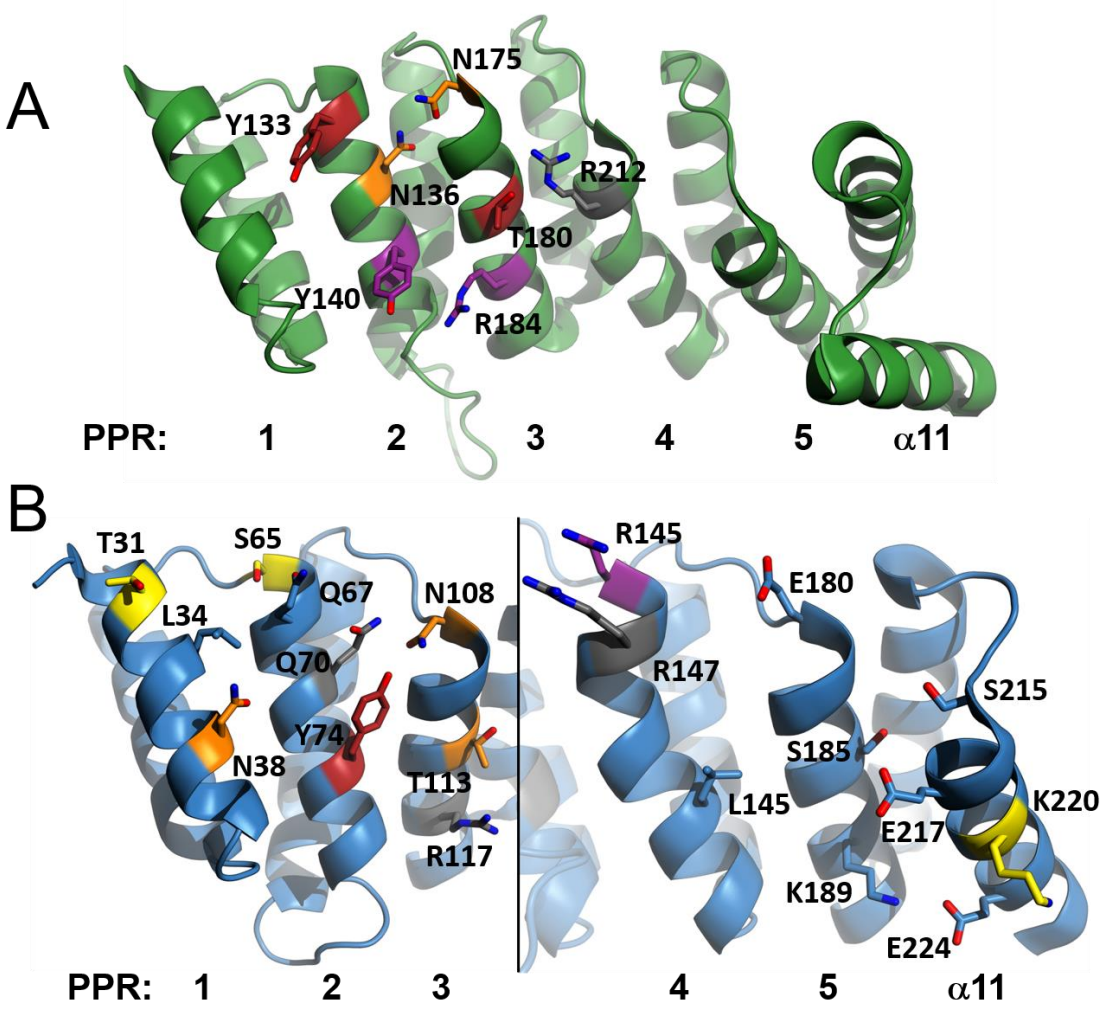


Figure 3-5: Residues selected for mutation in AtPRORP PPRs, generated in PyMol (16). (**A**) AtPRORP1 (PDB 4g24) PPRs are numbered left (PPR1) to right (helix α11). Residues that were targeted for mutation are numbered. Carbon atoms are color-coded by the largest effect on binding: 0.5–2-fold (blue), 2–3-fold (yellow), 3–10-fold (orange), 10–30-fold (red), and > 30-fold (purple). Arg 212 (gray) affinity could not be measured by anisotropy. (**B**) AtPRORP2 (PDB 5diz) PPRs are numbered left to right, 1-3 (left panel) and 4, 5, and helix α11 (right panel). Residues that were targeted for mutation are numbered. Carbon atoms are color-coded as in **A**, except that residues for which alanine mutants were not soluble are gray.

We measured the binding affinities of the *At*PRORP1 mutants for the *B. subtilis* 5-nt pre-tRNA^{Asp} substrate using the FA assay at 330 mM NaCl; these data are summarized in Table 3-2. We observed the largest reductions in binding affinity for the Y140A and R184A variants with decreases of 170- and 70-fold, respectively. Thr 180 was the only residue in a canonical PPR base-selection position that we

tested with a strong effect on binding: T180A reduces the binding affinity by approximately 40-fold compared to WT *At*PRORP1. The other canonical base-selecting residues mutated, N136A had a 4.5-fold effect, N175A had an approximately 8-fold effect. The final residue examined, R212A, eliminated binding as measured with the anisotropy assay (Fig. B-3B). Further, the WT level of enzymatic activity in an STO assay was not fully recovered with > 35 μM enzyme and high Mg^{2+} concentrations and included several significant miscleavage bands (Fig. B-3A).

Table 3-2. Na^+-dependence of *At*PRORP1 variants affinity for *B. subtilis* pre-$tRNA^{Asp}$.

*At*PRORP2 residue	*At*PRORP1 Variant	K_D (nM)[a]	Z[b]	log(K_0)[b]	ΔG_0 (kcal/mol)[c]
–	WT	160 ± 20	4.3 ± 0.3	5.0 ± 0.1	-6.9 ± 0.1
Q67	Y133A	2600 ± 500	4.4 ± 0.3	3.8 ± 0.2	-5.2 ± 0.3
	Y133F	4500 ± 800	4.2 ± 0.2	3.7 ± 0.1	-5.1 ± 0.1
Q70	N136A	1140 ± 50	3.9 ± 0.2	4.3 ± 0.1	-6.0 ± 0.1
Y74	Y140A	29600 ± 1200	4.2 ± 0.4	3.0 ± 0.3	-4.1 ± 0.4
	Y140F	1110 ± 40	4.1 ± 0.6	4.3 ± 0.3	-5.9 ± 0.4
N108	N175A	1400 ± 500	3.8 ± 0.2	4.4 ± 0.2	-5.9 ± 0.3
T113	T180A	3700 ± 800	4.0 ± 0.2	3.8 ± 0.1	-5.2 ± 0.1
	T180S	1830 ± 30	4.6 ± 0.4	3.9 ± 0.2	-5.4 ± 0.3
R117	R184A	10500 ± 2500	4.5 ± 0.5	3.2 ± 0.3	-4.4 ± 0.3
	R184K	1900 ± 400	4.6 ± 0.5	4.0 ± 0.3	-5.5 ± 0.4
R147	R212A[d]	> 300000	ND	ND	ND

[a] Mean and standard deviation reported are from two independent experiments in 330 mM NaCl.
[b] Value and error from fitting Equation 3-4 to the log-log plot of the Na^+-dependence data using $\varphi^{Na} = 0.88$ as described in the Materials and Methods (25).
[c] Calculated using $\Delta G_0 = -RT \times \ln K_0$.
[d] Affinity for the R212A mutant was not measureable; estimated from Fig. B-3B. Values from the Na^+-dependence plot were not determined (ND).

We parsed the determinants of substrate binding *At*PRORP1 in more detail by analyzing the salt dependence of the mutants. In general, the mutations had little effect on the Z value for the Na^+-dependence of binding affinity, but they affected the intercept value, K_0 (Table 3-2). These results indicate that the mutated side chains did not form ionic interactions with the phosphodiester backbone of pre-tRNA, rather forming non-ionic interactions with the substrate. The largest measurable reduction in affinity (170-fold) was observed for the Y140A variant, corresponding to a loss of 2.8 kcal/mol of non-ionic binding energy. However, the Y140F mutation only reduced K_0 by 5–6-fold, corresponding to a loss of

approximately 1 kcal/mol compared to WT *At*PRORP1. These results are consistent with PRORP using both the tyrosine hydroxyl and the phenyl ring to bind substrate (29). For the R184A mutant, the 70-fold reduced affinity corresponds to a reduction of 2.5 kcal/mol compared to WT, while the R184K mutation reduced the K_0 by 11-fold, corresponding to a loss of 1.4 kcal/mol compared to WT.

The Tyr 133 variants revealed a relationship different from that of the Tyr 140 variants. While the Y133F variant reduced affinity by 1.8 kcal/mol, the Y133A variant yielded a similar reduction at 1.6 kcal/mol. These data suggest the possibility that the hydroxyl group, but not the phenyl ring of Tyr 133, is contributing to substrate affinity. N136A and N175A mutations reduce the value of the intercept corresponding to approximately the loss of 1 kcal/mol apiece. We estimate a free energy loss for T180A of 1.7 kcal/mol from WT, while the T180S variant reduces the affinity to a similar extent, indicating a loss of 1.5 kcal/mol.

To determine whether additional PRORP PPR surfaces are important for substrate binding, we screened the pre-tRNA binding affinity of 22 alanine variants in *At*PRORP2 (Fig. 3-5B). Previously, we proposed that the 1st, 3rd, 6th and 10th residues (numbered as in Barkan, *et al*.) in each PPR motif could periodically contribute to substrate binding in *At*PRORPs (21, 26, 36, 37). Therefore, we targeted residues in these positions for all five PPR motifs and α11 for alanine mutagenesis. This analysis necessarily excludes three alanine residues (Ala 110, 150, and 182).

We examined the binding affinity of the alanine mutants with *A*. *thaliana* nuclear 6-nt pre-tRNA^{Gly} and *B*. *subtilis* 5-nt pre-tRNA^{Asp} substrates. The residues we identify with *At*PRORP2 mutagenesis are consistent with the surface identified in *At*PRORP1. The K_D values for pre-tRNA of the Q67A (position 3; PPR2), N108A (position 1; PPR3), T113A (position 6; PPR3), and R145A (position 1; PPR4) mutants increased by at least 1.5-fold compared to wild type *At*PRORP2 (Table B-2). This analysis also identified residues not identified in *At*PRORP1, though they fall primarily within the same surface. The K_D values for pre-tRNA of the N38A

(position 10; PPR1), S65A (position 1; PPR2), T31A (position 3; PPR1), and K220A (position 6; α11) mutants increased by at least 2-fold compared to wild type *At*PRORP2. Four *At*PRORP2 alanine mutants – Q70A (position 6; PPR2), Y74A (position 10; PPR2), R117A (position 10; PPR3), R147A (position 3; PPR4) – did not express as soluble proteins, suggesting that mutation of these residues may affect the stability of *At*PRORP2. These residues are all found within the proposed binding surface.

Given the importance of the equivalent *At*PRORP1 Tyr 140, we generated more conservative mutations (Y74S and F) of residue Y74 to investigate the type of interaction between this amino acid and the pre-tRNA. The Y74S lacks the ability to make stacking interactions. The Y74F variant maintains stacking interactions, yet lack the hydrogen-bonding capability of the tyrosyl hydroxyl group. We found that these mutations had significant impact on substrate binding (Table B-2). These results indicate that this residue, like Tyr 140 in *At*PRORP1, interacts with substrate using both the phenyl ring and the hydroxyl group. Taken together, the above results indicate that surface we identified in *At*PRORP1, primarily in PPR2 and PPR3, is the major surface involved in PRORP substrate binding.

To investigate whether *At*PRORPs recognize pre-tRNA in a base-specific manner, we analyzed the effects of several mutations in pre-tRNA^{Asp}. *At*PRORP1 contains a binding site between PPR motifs 2 and 3 (Y133/N136/N175) that should recognize pyrimidines based on previously established PPR recognition codes (21, 22). The D- and TψC-loops (tRNA elbow) in pre-tRNA have been proposed to interact with the PPR domain (18). For this analysis, we examined pyrimidines that would likely interact with PRORPs. We assumed that the pyrimidines should not be in secondary/tertiary contacts or otherwise buried and inaccessible in the unbound tRNA structure. This limited us to the uridines at positions 16, 17, 20, and 21 (Fig. 3-2A). Mutations at these positions to adenosine altered the affinity by at most 2-fold (Table B-4). In combination with the previous data regarding the effects of tRNA mutations on PRORP binding/catalysis, these data reinforce the hypothesis that PRORPs utilize a mode of substrate recognition different from the

previously described PPR base-selection (14, 19). However, it may still be that there are sites that interact with PRORP PPRs in sequence-specific manner and that we have yet to identify them.

Na^+ screening inhibits *At*PRORP1 single-turnover activity

While it is possible that the ionic strength affects only the binding affinity, it might also affect other aspects of PRORP catalysis. To determine whether the NaCl concentration affects cleavage catalyzed by PRORP, we performed single-turnover (STO) activity assays with 5 μM *At*PRORP1, which is saturating under low-NaCl conditions, and limiting *B*. *subtilis* pre-$tRNA^{Asp}$. We used concentrations of $MgCl_2$ that we previously determined to be either saturating (20 mM) or sub-saturating (1.25 mM) for catalysis (12). The observed rate constant (k_{obs}) is independent of the NaCl concentration below ≈ 200 mM, but is reduced at higher concentrations (Fig. 3-6). The concentration dependence of NaCl inhibition above 400 mM is similar for saturating and sub-saturating $MgCl_2$. Fitting a general inhibition model (Equation 3-3, weighted fit, Materials and Methods) with a variable Hill Coefficient (n^{Na}) to the data yields similar IC_{50} (310 ± 70 mM for 20 mM $MgCl_2$ and 360 ± 70 mM for 1.25 mM $MgCl_2$) and n^{Na} values (4.5 ± 1.3 for 20 mM $MgCl_2$ and 5.0 ± 1.3 for 1.25 mM $MgCl_2$) for both $MgCl_2$ conditions (Fig. 3-6). The Hill coefficients for STO inhibition (4.5–5) are in good agreement with the cooperativity we observe for binding inhibition. The enzyme concentration is saturating (more than 5-times the K_D, as measured in Ca^{2+}) until NaCl ≥ 530 mM, yet the activity is inhibited at lower concentrations. If the binding affinity is equivalent in Ca^{2+} and Mg^{2+}, the data suggest that NaCl affects kinetic steps in addition to binding.

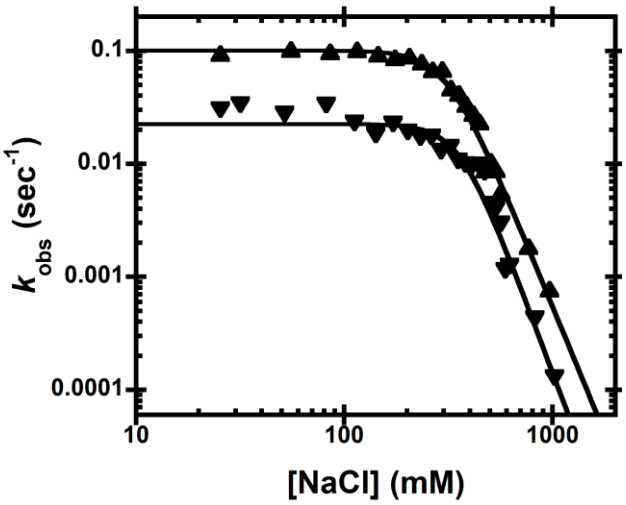


Figure 3-6: Na^+-dependence of AtPRORP1 catalysis. AtPRORP1 catalysis under single turnover conditions with pre-$tRNA^{Asp}$, plotted as the k_{obs} versus [NaCl]. Data include AtPRORP1 catalysis in 20 mM (▲) or 1.25 mM (▼) $MgCl_2$. Equation 3-3 (Materials and Methods, weighted fit) was fit to the data. For 20 mM $MgCl_2$, IC_{50} = 310 ± 70, n^{Na} = 4.5 ± 1.3. For 1.25 mM $MgCl_2$, IC_{50} = 360 ± 70, n^{Na} = 5.0 ± 1.3.

DISCUSSION

The goal of this work was to characterize the molecular interactions *At*PRORPs make with their substrates. We used the ion-dependence of binding and catalysis and mutations in the PPR domain to characterize the *At*PRORPs-substrate complex. The NaCl-dependence revealed 4–5 interactions with phosphodiester bonds in pre-tRNA. Comparison of Na-dependence and mutation of residues in the substrate binding domain in both *At*PRORP1 and 2 reinforces that there are differences in substrate recognition, despite having well-conserved structural and functional features (27). Our extensive mutagenesis data in the PPR domain allow us to contrast PRORP binding with previously-described PPR proteins.

PRORP PPR domain

PPR-containing proteins are a large family with the structurally-conserved ≈ 35 residue helix-turn-helix motif found in tandem repeats that have been implicated in RNA metabolism (38, 39). The number of repeats vary from as few as 2 repeats in mitochondrial RNA polymerase to as many as 30 repeats in some plant proteins (39, 40). PPR proteins are found broadly in eukaryotes, with land plants having the largest set of PPR proteins. For example, there are over 400 members in

Arabidopsis and *Oryza* (41). Some PPR proteins bind target RNAs in a sequence-specific manner, with recognition of a nucleobase achieved primarily by residues 6 and 1' (also numbered 4 and ii by Yagi, *et al*. (22), or 5 and 35 by Yin, *et al*. (37)) on the A helices of two tandem PPR motifs (Fig. 3-1B, A helices colored in dark green) (21). In some cases, the binding sites have been identified in the UTRs of mRNAs, where the PPR proteins are proposed to regulate the splicing, translation and/or stability of the mature transcript (39, 42). One goal of this work was to determine whether PRORP PPR domains use the same or similar mechanism to recognize pre-tRNAs.

PRORP PPRs are frequently non-canonical at base-specifying residues when compared to ssRNA-binding PPR proteins (e.g., 6 and 1' are not often Asn, Asp, or Thr; Fig. B-1). Thus, we hypothesized that they are unlikely to utilize the same mode of base selection by tandem repeats. For instance, conservation to other PPR domains yielded predictions of only two or three PPR motifs in PRORPs (6, 7), yet crystal structures revealed five tandem repeats in *At*PRORP1 (Fig. 3-1A) (13). By homology, the metazoan PRORP also contains at least five repeats, the last three of which have been visualized in a crystal structure of human PRORP with an N-terminal deletion (43). Several substrate binding residues that we identified, such as Tyr 140 and Arg 184, are conserved in metazoan PRORPs (Tyr 183 and Arg 218 in humans). Thus, even though the metazoan PRORPs require additional subunits for catalysis, pre-tRNA recognition by the PPR domain may utilize the same surface as plant PRORP PPRs.

The PPR residue for which alanine mutation was most detrimental to substrate affinity was Tyr 140 in *At*PRORP1 PPR2 (170-fold). The equivalent mutant in *At*PRORP2 (Y74A) was insoluble, but Y74S reduced the binding affinity 23-fold. The second largest effect was from the alanine mutant of Arg 184 in *At*PRORP1 PPR3 (70-fold). The equivalent R117A in *At*PRORP2 was insoluble. Each of the effects for the *At*PRORP1 mutations are greater than the 34-fold decrease reported for a Δ245 *At*PRORP1, which fully lacks the first 4 PPR motifs (13). This effect observed with the PPR deletion was measured at 1 mM $CaCl_2$ and 100 mM

NaCl with the *A*. *thaliana* mitochondrial pre-tRNA^{Cys} substrate. Under these conditions, the Y140A and R184A mutations have less effect, with affinities up to 25% tighter than WT (Table B-3). Like the *B*. *subtilis* pre-tRNA^{Asp} substrate, Y140A and R184A bind the *A*. *thaliana* pre-tRNA^{Cys} substrate weaker than WT under the 20 mM $CaCl_2$ and 330 mM NaCl conditions (> 93-fold and > 3.8-fold, respectively). This effect can be explained by the ionic portion of the substrate affinity, which adds a significant portion of the PRORP-substrate affinity under lower ionic strength. Thus, individual non-ionic interactions contribute less to the overall affinity under lower NaCl concentrations.

The loss of 1 kcal/mol that we observe for the Y140F is consistent with Tyr 140 interacting through a hydrogen bond to substrate with the tyrosine hydroxyl. The additional 1.8 kcal/mol loss observed with the alanine substitution is consistent with the energies from stacking a phenyl ring with a nucleic acid base (29). For the R184A mutant, the 70-fold reduced affinity corresponds to a reduction of 2.5 kcal/mol compared to WT, which could indicate loss of as many as 2–3 hydrogen bonds from the guanidinium group, fewer hydrogen bonds due to the charged nature of the side chain, and/or hydrophobic interactions with the arginine methylene groups. By comparison, the R184K mutation reduced the K_0 by 11-fold, corresponding to a loss of 1.4 kcal/mol compared to WT, consistent with the loss of a hydrogen bond. In *At*PRORP1, Tyr 140 and Arg 184 are located at position 10 of neighboring helices forming a structural, non-sequential, YR pair. The YR pair we identified as important for substrate binding is widely conserved in PRORP PPRs, including metazoan PRORPs (Fig. B-2).

Bioinformatic studies have demonstrated that both Arg and Tyr are enriched in RNA-binding sites on proteins (44). However, to our knowledge the combination YR pair has not been identified more generally as a significant RNA binding motif. That said, PRORPs are not the first instance of a YR pair identified on the interaction surface of an RNA binding protein. The C-terminal La domain of the telomerase protein p65 contains a highly conserved YR pair (Tyr 407 and Arg 465) situated on neighboring β-strands and are important for recognition of the

conserved GA bulge in stem IV of the telomerase RNA (45). Given the significant contribution of the phenyl ring and the guanidinium group revealed by the Y140A/F and R184A/K mutations in *At*PRORP1 and Y74S/F in AtPRORP2, we propose that these residues make similar interactions with the tRNA elbow, the conserved structural feature that results from the interaction of the D- and TψC-loops. These interactions do not have to be sequence specific, but like the p65 YR pair (45), they could favor purines such as the conserved $G_{18}G_{19}$ in the D-loop.

In contrast to the Tyr 140 data, the data for Tyr 133 indicate that the hydroxyl group, but not the phenyl ring, contributes to substrate affinity. The data for N136A and N175A correspond to the loss of approximately 1 kcal/mol apiece, consistent with a hydrogen bond from the amide side chain of each. The data for T180A indicate 1.7 kcal/mol loss, consistent with up to one or more hydrogen bonds and/or hydrophobic interactions. The T180S variant reduces the affinity to a similar extent, indicating a loss of 1.5 kcal/mol, which may indicate that the Thr methyl group is required either to sterically position the hydroxyl for substrate interaction or by a hydrophobic contact to substrate.

While *At*PRORP1 Thr 180 (Thr 113 in *At*PRORP2) is at a base-selecting position 6, its corresponding 1’ partner in the putative base-selection position would be Arg 210 (Arg 145 in *At*PRORP2). Arginine residues have not previously been identified as base-selecting residues for PPRs (21, 22). Furthermore, the potential nucleobase binding cleft is occluded by an interaction between Thr 180/133 and Arg 212/147 in the *At*PRORP1 and *At*PRORP2 crystal structures (PDB 4g24 and 5diz) (13, 26). The R212A mutant showed severe defects in substrate binding and catalysis, though the mutation may disrupt pre-tRNA binding by altering the structure of the PPRs. Arg 212 is involved in multiple interactions within an extended series of salt-bridge and hydrogen bonding interactions that begins at Thr 180 in helix A of PPR3 and terminates in helix B of PPR 4 at Ser 240 (Fig. B-3C) (13). While we could not express the equivalent *At*PRORP2 R147A as soluble protein, we tested the *At*PRORP1 R212A stability by the ThermoFluor assay. However, the observed melting point was unchanged from WT (data not

shown). Taken together, the mutagenesis data for *At*PRORP1 and *At*PRORP2 indicate that the PRORP PPR domain does not interact with pre-tRNA in the same manner as ssRNA-binding PPR proteins.

An alternative recognition mode could include structural recognition, like the tRNA elbow recognition employed by the specificity (S) domain of the P RNA in the bacterial ribozyme. The S-domain makes stacking interactions with the conserved G_{19}–C_{56} tertiary interaction between the D- and TψC-loops and sugar face interactions with the conserved G_{53}–C_{61} pair at the end of the TψC arm (46). These non-specific interactions with pre-tRNA recognize conserved structural elements and add to the ability of P RNA to recognize the entire set of pre-tRNA transcripts. Likewise, we propose that the PPRs of PRORPs recognize tRNA structure using a similar strategy through structure-specific interactions. Nonetheless, while our data support a model in which PRORP PPRs recognize the tRNA structure rather than conserved sequences, we cannot rule out at least the partial involvement of base-recognition strategies as those employed by canonical PPR proteins.

In further support of substrates selection by PRORPs on a structural basis, the base-selecting residues conserved in other PPR proteins have low conservation in PRORPs. For *At*PRORP1, the residues at potentially base-selecting sites which we excluded include proline 96, lysine 101, glycine 215, and alanine 98 and 177. These residues were not previously identified as base-selecting and are not conserved within PRORP PPRs. We also excluded serine 95 and 131 (in *At*PRORP1), which have the potential for canonical PPR base-selection. However, these residues do not have a base-selecting residue at on the neighboring PPR motif and are not conserved. Ser 95 is found as Ala in other PRORP homologues and Ser 131 frequently found as Asn. While these differences could reflect changes in substrate specificity between homologous PRORPs, it is more likely that these positions are not involved in base selection given that PRORPs must recognize similar tRNA structures across various eukaryotes.

The PRORPs from the moss *Physcomitrella patens* appear to have a total of six PPR motifs (Fig. B-1) (47, 48). The sequence aligns to the extended loop in *Arabidopsis* PRORP PPRs (Fig. 3-1B, Fig. B-1). The additional helices generate new PPR binding clefts consistent with established PPR nucleobase recognition motifs (N288/D312, S318/N340, and E345/N375). These binding clefts are different from those observed in *Arabidopsis* PRORPs. Given the presence of additional base-selecting sites, *Physcomitrella* PRORPs might rely more strongly on base-selection for pre-tRNA recognition, or recognize alternative substrates. All three *Physcomitrella* PRORPs contain the additional repeat, while all three *Arabidopsis* PRORPs and metazoan PRORPs lack it, suggesting that the PPR motif insertion occurred after the divergence of mosses from other land plants. Thus, PRORPs from certain clades may utilize variations on the recognition strategy we have defined.

PRORP-substrate recognition model

Using the data we have presented here, as well as previously published data (14, 18), we propose a model to describe how *At*PRORP1 and *At*PRORP2 may bind their substrates. There are no crystal structures available for any of the tRNAs used in this study to use for our PRORP-tRNA model. However, since the 3D structure of tRNAs is highly conserved, we have chosen a tRNA that resembles a canonical eukaryotic tRNA with an available crystal structure (PDB 1ehz) (49). For constructing the model, we have employed the recently solved crystal structures of *At*PRORP1 (PDB 4g24) and *At*PRORP2 (PDB 5diz) (13, 26). The two structures are highly conserved, yet there is a significant difference in the angle between specific PPR motifs and between the central domain and the metallonuclease domain; as a consequence, *At*PRORP2 is found in a more “open” conformation. We proposed that the two distinct structural snapshots may represent two different conformations that potentially play a role in substrate binding (26). Moreover, we found that the more “open” *At*PRORP2 conformation can more readily accommodate tRNA, since the distance between the active site and the proposed substrate binding site in the PPR is ~ 50 Å. Therefore, we used the tRNA-

*At*PRORP2 interaction model for generating a model for *At*PRORP1-tRNA recognition. A previous attempt to model the complex utilized the *At*PRORP1 bound to tRNA with base-specifying interactions between a TψC-loop cytosine and PPR motifs 3 and 4. To accomplish this, the TψC-loop coordinates were modified using a tRNA bound to a pseudouridine synthetase (50), with the remaining tRNA coordinates derived from the bacterial RNase P holoenzyme in complex with a tRNA product (46).

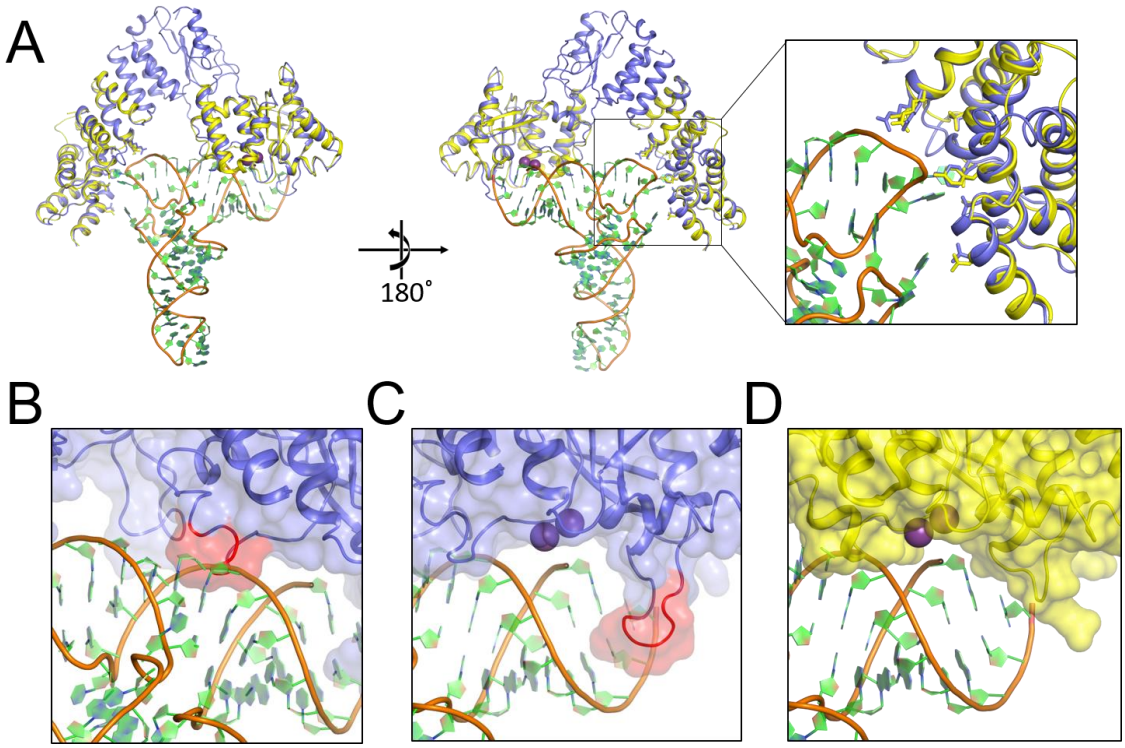


Figure 3-7: Model of the PRORP-substrate complex. (**A**) Overall view of the modeled complex. AtPRORP2 (PDB 5diz, blue) is bound to tRNA (PDB 1ehz). The AtPRORP1 (PDB 4g24, yellow) NYN and PPR domains are aligned to the corresponding residues domains in AtPRORP2. Close-up view of the PPR domain highlights the positions of the residues that we identified for which mutation affected binding. (**B**) Close-up of the AtPRORP2-tRNA complex with a potential steric clash between NYN helix α21 (red) and the 3' side of the tRNA acceptor stem. (**C-D**) Close-up of the AtPRORP2- (C) or AtPRORP1-substrate complex (D) showing the NYN active site. The loop for which AtPRORP2–3 have a 4 residue insertion is highlighted in red (C). In both structures, the 5' leader of pre-tRNA would extend forward from the panel, while the 3' trailer would extend behind the NYN domain.

The tRNA substrate can be accommodated relatively well between the metallonuclease and PPR domains in our model (Fig. 3-7), with one exception. There is a short helix (α21) and part of the loop that precedes it near the active site that sterically clashes with the 3' strand of the acceptor stem. We posit that

this helix will adopt a different conformation upon pre-tRNA binding and may be directly involved in recognition. This region could serve as a hinge that allows or blocks substrate binding to the metallonuclease domain. Interestingly, the invariant and solvent exposed Arg 496 (Arg 443 in PRORP2) and His 498 (His 445 in PRORP2) residues that reside in this region would be ideally placed for interaction with the phosphodiester backbone in the acceptor stem. We previously mutated the α21 residue His 498 (H498A/H498Q, *At*PRORP1) or His 445 (H445A, *At*PRORP2); all mutations reduced the STO k_{obs} without significantly affecting the K_D (12, 26). We proposed this residue was involved in positioning the substrate and our model provides potential contacts with substrate for testing this hypothesis.

Our model suggests an exit groove for the 5' leader that would place the N_{-3}/N_{-2} phosphodiester bond and N_{-3} nucleoside outside the bounds of the NYN domain. This is consistent with our Na^+-dependence data for tRNA and 1-nt pre-tRNA, which suggested few contacts beyond the N_{-2} nucleoside. Moreover, our model places the N_{-2}/N_{-1} phosphodiester bond near the invariant His 438 (386 in PRORP2) and Arg 441 (389 in PRORP2), implicating these residues for interacting with the negatively charged backbone. This aspect of our model appears to be congruent with our data indicating that PRORPs interact with the tRNA leader primarily through one non–base-specific phosphodiester backbone contact. Furthermore, the α16–α17 loop is positioned to separate the 5' leader and 3' trailer. Interestingly, *At*PRORP1 has 4 fewer amino acids in the loop than *At*PRORP2/3. These loop differences might explain the variations in 5' end discrimination we observed previously, in which *At*PRORP2/3 had a stronger propensity to miscleave at the N_{-1} position when an N_{-1}/N_{73} base pair was possible (20). Thus, there may be differences in the 5' end recognition employed by the two loops.

The elbow region of tRNAs are highly structured and numerous tRNA binding enzymes recognize this part of the tRNA using a variety of interaction motifs (51). Importantly, our model predicts that i) Arg 210 (Arg 145 in *At*PRORP2) and Arg 212 (Arg 147 in PRORP2) would be involved in contacting TψC stem

phosphodiester bonds, ii) Asp 105 (Asn 38 in PRORP2), Asn 136 (Gln 70 in PRORP2), and Asn 175 (Asn 108 in PRORP2) are positioned to hydrogen bond with bases of the D loop (first residue) or the TψC (last two) loop respectively, and iii) Tyr 140 (Tyr 74 in PRORP2) is capable of both base stacking and hydrogen bond interacting with bases of the TψC loop. All these model-based observations are consistent with the data presented herein. Although our proposed model needs to be further tested, it provides insight into the details of precursor tRNA binding of PRORPs in the absence of crystal structures of PRORP-tRNA complexes and a rough framework for the design of future experiments.

PRORP kinetic mechanism

In addition to decreasing pre-tRNA affinity, the NaCl concentration also inhibits pre-tRNA cleavage catalyzed by *At*PRORP1 under single turnover conditions. We previously proposed a minimal kinetic mechanism that included only binding and chemistry steps (Scheme 3-1) (12). For this mechanism, if inhibition by NaCl is only caused by decreasing the concentration of the enzyme-substrate complex, then IC_{50} should correlate with the transition from enzyme at saturation to sub-saturation. However, the IC_{50} corresponds to NaCl concentrations when the enzyme concentration is saturating in $CaCl_2$ ([E] > 20*K_D) (Table 3-1, Fig. 3-6). It remains possible that the $K_{1/2}$ in Mg^{2+} is significantly weaker than the K_D in Ca^{2+}. Assuming the Mg^{2+} $K_{1/2}$ and Ca^{2+} K_D are similar, one possible explanation for this result is that NaCl inhibits cleavage, possibly by Cl^- replacing the metal-water ligand as occurs in Zn^{2+}-dependent enzymes (52). Alternatively, this effect could be due to a more complicated mechanism.

$$E + S \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} ES \xrightarrow{k_{chem}} EP$$

Scheme 3-1: Minimal kinetic mechanism of PRORP (E) binding to pre-tRNA (S) to yield 5' matured tRNA (P).

The reduced binding affinity we observe for higher NaCl concentrations could result from changes in either or both k_{on} and k_{off}. That is, when the NaCl concentration increases, then k_{off} increases, k_{on} decreases, or both. In these

cases, increasing NaCl concentration increases the $K_{1/2}$ ((k_{off}+k_{chem})/k_{on}) until the enzyme is no longer saturating and the k_{obs} decreases. For LacI, k_{on} decreases with the NaCl concentration above 100 mM, while k_{off} increases with the NaCl concentration (53). We simulated this behavior assuming that NaCl increases the k_{off} or does not affect it, while the k_{on} decreases by NaCl. From these simulations, the observe decrease in k_{obs} requires $k_{off} < k_{chem}$ whether k_{off} is affected by NaCl or not. Thus, if NaCl only affects the binding affinity in a minimal kinetic mechanism, our NaCl inhibition experiments require that *At*PRORP1 does not bind pre-tRNA in rapid equilibrium.

Alternatively, if the binding event is rapid equilibrium, then the data require that additional kinetic steps of PRORP catalysis are affected by NaCl. One possibility is that NaCl affects the active site, such as by Cl^- displacing the metal-water nucleophile (52). For this hypothesis, k_{obs} should decrease in a manner that does not correlate directly to the increase in K_D. However, the Hill coefficient for Na^+ kinetic inhibition was similar to the cooperativity of the binding inhibition. This suggests that inhibition by NaCl is primarily due to the enzyme-substrate complex binding affinity or conformation. For instance, plant PRORPs might utilize a substrate recognition mechanism similar to the bacterial RNase P, including a kinetic step after binding such as a conformation change (Scheme 3-2) (54, 55). If the second step requires displacement of Na^+ ions in addition to those displaced by the initial binding event, then k_{obs} could be reduced at NaCl concentrations in which the enzyme concentration would otherwise be saturating. More detailed studies of *At*PRORP binding kinetics – and how they are affected by NaCl – are needed to distinguish these hypotheses and develop a more accurate kinetic mechanism for *At*PRORPs.

$$E + S \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} ES \underset{k_{-1}}{\overset{k_{1}}{\rightleftharpoons}} ES^* \xrightarrow{k_{chem}} EP$$

Scheme 3-2: Kinetic mechanism including transition to an active complex (ES*).

CONCLUSIONS

The data we present herein support a novel model for PRORP-substrate recognition that shares similarities with the mode of substrate recognition by the RNase P ribozyme. To begin with, we characterized the salt-dependence of PRORP-substrate binding to parse the ionic and non-ionic contributions to PRORP-substrate binding affinity. The data revealed that *At*PRORPs make at least four contacts with pre-tRNA phosphodiester bonds. Furthermore, only one of these is contained in the leader sequence, most likely at the N_{-2}/N_{-1} phosphodiester bond. Additionally, we identified an extended surface on the PPR domains of *At*PRORP1 and *At*PRORP2 that interact with substrate. Mutations on this surface suggested a mode of binding similar to that of the RNase P ribozyme, yet different from that of sequence specific ssRNA-binding PPR proteins. The biochemical and modeling data we have presented will facilitate the development of additional hypotheses for single-subunit PRORP substrate recognition. Given that two additional proteins are required for catalysis, there are likely differences for metazoan PRORPs that will need to be determined.

MATERIALS AND METHODS

Reagents

Reagents used in this chapter include 3-morpholinopropane-1-sulfonic acid (MOPS; Acros Organics), tris(2-carboxyethyl)phosphine (TCEP; Gold Biotechnology, Inc.), 1,4-dithio-D-threitol (DTT; Gold Biotechnology, Inc.) 5-iodoacetamidofluorescein (5-IAF; Life Technologies), ethylenediaminetetraacetic acid (EDTA; Acros Organics), tris(hydroxymethyl)aminomethane (Tris; Fisher Scientific), magnesium chloride ($MgCl_2$; Fisher Scientific), calcium chloride ($CaCl_2$; Sigma-Aldrich), ammonium chloride (NH_4Cl; Fisher Scientific), lithium chloride (LiCl; Fisher Scientific), potassium chloride (KCl; Fisher Scientific), rubidium chloride (RbCl; Sigma-Aldrich), sodium chloride (NaCl; Sigma-Aldrich), ammonium acetate (NaAc; Fisher Scientific), sodium bromide (NaBr; Sigma-Aldrich), sodium iodide (NaI; Sigma-Aldrich), sodium nitrate ($NaNO_3$; Sigma-Aldrich), sodium

sulfate (Na_2SO_4; Fisher Scientific), 5'-O-monophosphorothioate guanosine (GMPS, synthesized), nucleoside triphosphates (NTPs; Sigma-Aldrich), kanamycin (Acros Organics), chloramphenicol (Acros Organics), urea (MP Biochemicals), bromophenol blue (BPB; Fisher Scientific), xylene cyanol (XC; United States Biochemical Corporation), bulk yeast tRNA (Fisher Scientific), and the SequaGel UreaGel system (National Diagnostics).

Enzyme preparation

Wild type Δ76 *At*PRORP1 and full length *At*PRORP2 enzymes were expressed and purified as described previously (20). Variants were generated by site directed mutagenesis using *PfuTurbo* DNA Polymerase (Agilent Technologies) and DNA oligo primers (IDT) (56). Sequences were verified by the Sanger method at the University of Michigan DNA Sequencing Core facility. Variants were expressed in Rosetta™, Rosetta 2™ or BL21(DE3) *E. coli* (Novagen/EMD Millipore) from the T7 promoter on a pETM-11 (encoding His_6-TEV-*At*PRORP1) or pMCSG7 (His_6-TEV-*At*PRORP2) vector in LB media with 50 µg/mL kanamycin and 33 µg/mL chloramphenicol for selection of pETM-11 and pRARE (a plasmid encoding rare-codon tRNAs in the Rosetta™ cell lines) and 100 µg/mL ampicillin for selection of pMCSG7, respectively. All variants were purified as described previously (12, 13), including removal of the N-terminal His_6-tag by TEV protease, treatment with ethylenediaminetetraacetic acid (EDTA) to remove metal ions, flash freezing with liquid N_2, and storage at -80°C in 20 mM 3-morpholinopropane-1-sulfonic acid (MOPS) pH 7.8, 100 mM NaCl, and 1 mM tris(2-carboxyethyl)phosphine (TCEP, *At*PRORP1) or 1 mM 1,4-dithio-D-threitol (DTT, *At*PRORP2).

Substrate preparation

Substrates were prepared as described previously (12, 13). Briefly, substrates were synthesized by run-off transcription from restriction-digested plasmid encoding pre-tRNA, a PCR-amplified template DNA, or a commercially synthetized ultramer oligo (IDT) (57). *In vitro* transcription was initiated with 5'-O-monophosphorothioate guanosine (GMPS) in 5:1 excess of GTP, then the products reacted with 5-iodoacetamidofluorescein (5-IAF) to generate a 5'-

fluorescein label. The correct pre-tRNA products were then gel purified by 12% urea-PAGE and substrates eluted from the gel using the crush-soak method (57). The purified pre-tRNAs were washed and concentrated using 10 kDa MWCO Amicon® Ultra Centrifugal Filters, then ethanol precipitated. Substrate stocks were resuspended in 10 mM tris(hydroxymethyl)aminomethane (Tris) pH 8.0 with 1 mM EDTA and quantified at 260 nm. The extinction coefficient for *Bacillus subtilis* pre-tRNA^{Asp} is 685000 M^{-1} cm^{-1} (experimentally determined by alkaline hydrolysis), for *A*. *thaliana* mitochondrial pre-tRNA^{Cys} is 674390 M^{-1} cm^{-1} (experimentally determined by alkaline hydrolysis), and *A*. *thaliana* nuclear pre-tRNA^{Gly} 870700 M^{-1} cm^{-1} (calculated) for total RNA concentration and 492 nm (extinction coefficient = 78000 M^{-1} cm^{-1}) for fluorescein concentration, then stored at -20 or -80°C. Immediately prior to an assay substrates were thawed, diluted with H_2O, and heated at 95°C for 60–90 seconds. Substrates were re-folded by cooling to 25°C for ≥ 10 minutes, then incubating with buffer (as specified for each assay) for ≥ 10 minutes.

Anisotropy binding assays

Thermodynamic binding assays were performed in a 96-well plate format as previously described (12, 13). Briefly, WT *At*PRORP1 was serially diluted and mixed 1:1 with low concentrations of a pre-folded pre-tRNA with variable leader lengths and a 5-fluorescein. In all experiments reported, a hyperbolic binding curve (Equation 3-1) was observed, with a maximum enzyme concentration ([P]) at least three times greater than the K_D and pre-tRNA concentration at least five times lower than the K_D. Reactions were incubated at 28 ± 1°C in 30 mM MOPS pH 7.8, and 1 mM TCEP (*At*PRORP1) or 1 mM DTT (*At*PRORP2). NaCl is variable as indicated for a given assay. Unless otherwise specified, we used 20 or 6 mM $CaCl_2$ for *At*PRORP1 and *At*PRORP2, respectively. Anisotropy of the 5'-fluorescein was measured with a Tecan Ultra plate reader with polarizing filters using excitation and emission wavelengths of 485 and 535 nm, respectively. Readings were taken 3–5 times over the course of 15–20 minutes to ensure there was no time-dependence in the data.

$$FA = FA_0 + \frac{\Delta FA \bullet [P]}{[P] + K_D} \quad (3\text{-}1)$$

Single-turnover assays

Single-turnover kinetic assays for *At*PRORP1 were performed in a stopped-format as previously described (12, 13). Briefly, enzyme was mixed with pre-folded *B. subtilis* pre-tRNA^{Asp} with a 5-nt leader and a 5'-fluorescein to final concentrations of 5 μM and 30 nM, respectively. Reactions were incubated at 25°C in 30 mM MOPS pH 7.8, 1 mM TCEP, with $MgCl_2$ and NaCl varied as indicated for a given assay. Aliquots were removed at various times and mixed 1:1 with a 2x quench dye (6 M urea (MP Biochemicals), 100 mM EDTA (Acros Organics), 0.1% bromophenol blue (BPB; Fisher Scientific), 0.1% xylene cyanol (XC; United States Biochemical Corporation), and 2 μg/μL bulk yeast tRNA (Fisher Scientific)). Products were resolved from substrate by ≥ 20% urea-PAGE and the gels were scanned using a Typhoon 9410 (GE Life Sciences) in fluorescence mode with a 532 nm green laser and fluorescein emission filter. Assays for *At*PRORP2 were carried out using the same conditions, but changes in polarization upon cleavage were detected using a ClarioStar in a 96-well plate format. The observed rate constants (k_{obs}) reported were determined by quantifying the fraction product using ImageQuant 5.2 software and fitting a single exponential (Equation 3-2, where *A* is the endpoint, *B* is the amplitude, and *t* is the time) to the data using KaleidaGraph 4.0 software. At low concentrations of NaCl (below ≈ 90 mM), the 5' leader product degraded after it appeared and did not accumulate to 100%. These data were fit to a double exponential and the k_{obs} from the first phase is reported. *At*PRORP1 inhibition by NaCl was determined by plotting the STO k_{obs} against the NaCl concentration and fitting Equation 3-3 to the data (as described in the Results section).

$$\text{Fraction Product} = A - B(e^{-k_{obs} * t}) \quad (3\text{-}2)$$

$$k_{obs} = \frac{k_{max}}{\left(1 + \left(\frac{[Na^+]}{IC_{50}}\right)^n\right)} \quad (3\text{-}3)$$

Cation/anion variation

We varied the cation and anion identities as XCl and Na_nX^{n-}, respectively. Thermodynamic binding assays were performed in a 96-well plate format as described above. Reactions were incubated at 28 ± 1°C in 30 mM MOPS pH 7.8, 1 mM TCEP, with 20 mM $CaCl_2$ and 330 mM of each cation, as listed for a given assay. The salts used were NaCl, NH_4Cl, RbCl, KCl, LiCl, NaAc, NaBr, NaI, $NaNO_3$, and Na_2SO_4. The data are reported as the mean and standard deviation from two independent experiments.

Sodium dependence

Equation 3-4 approximates Equation B-1 for cases in which effects on affinity from pH, anions, and divalent ions are negligible or can otherwise be precluded. The dependent variable is the monovalent cation concentration ([M^+]). The parameters include a "standard affinity" at 1 M M^+ (K_0), the number of phosphodiester bonds on the substrate interacting with the protein (Z), and the fraction of phosphodiester bonds in the nucleic acid that thermodynamically associate with a monovalent ion (φ). When varying divalent cations in the absence of monovalent ions, the slope is distinguished by replacing φ with ϕ.

$$-\log K_D = \log K_0 - Z\varphi \bullet \log[M^+] \quad (3\text{-}4)$$

Model building

Crystal structures of *At*PRORP1 (PDB ID: 4g23) and *At*PRORP2 (PDB ID: 5diz) and yeast $tRNA^{Phe}$ (PDB ID: 1ehz) were used to fit the elbow region to the proteins. Initial models were obtained using ZDOCK server (58) and these were processed through iterative rounds of manual adjustment by PyMOL (16). The model amino acid or nucleotide geometry regularization and use of allowed side chain rotomers were corrected with Coot (59). The coordinates of these models are available upon request.

APPENDIX B

This appendix contains supporting methods, tables, and figures for Chapter 3.

Methods

Protein-nucleic acid interactions

Equation B-1 was adapted from Equation 18 of deHaseth, *et al*. 1977 and describes the effects of solution ions on the binding affinity between proteins and nucleic acids (25). Variables affecting the slope include pH ($K_H[H^+]$), anion concentration ($K_X[X^-]$), monovalent cation concentration ($[M^+]$), and divalent cation concentration ($[S]/[S_0]$). Parameters include a "standard affinity" under 1 M M^+ (K_0), the number of proton donating groups on the protein directly interacting with substrate phosphates (r), the number of specifically-occupied anion binding sites on the protein (a), the number of phosphates on the substrate directly interacting with the protein (Z), the fraction of phosphates in the nucleic acid that thermodynamically associate with a monovalent ion (φ), and the competition of M^+ with M^{2+} for occupancy on backbone phosphates given by the total concentration of nucleotides $[S]$ and the concentration of nucleotides that are bound as they would with no M^{2+} in solution $[S_0]$. The term $[S]/[S_0]$ is described by Equation B-2.

$$-log\, K_{\mathrm{D}} = log\, K_0 - r \bullet log\left(\frac{1 + K_H[H^+]}{K_H[H^+]}\right) - a \bullet log(1 + K_X[X^-]) - Z\varphi \bullet log[M^+] - Z \bullet log\left(\frac{[S]}{[S_0]}\right) \quad \text{(B-1)}$$

$$\frac{[S]}{[S_0]} = \frac{1}{2}\left(1 + \sqrt{1 + 4K_A^M[M^{2+}]}\right) \quad \text{(B-2)}$$

Tables

Table B-1: Effects of cation and anion identity on affinity for pre-tRNA^{Asp}.[a]

Cation[b]	K_D (nM)[c]	Anion[d]	K_D (nM)[c]
Li^+	60 ± 10	$CH_3CO_2^-$	33 ± 6
K^+	72 ± 2	SO_4^{2-}	80 ± 30
Rb^+	70 ± 20	Cl^-	160 ± 20
NH_4^+	80 ± 60	NO_3^-	490 ± 60
Na^+	160 ± 20	Br^-	560 ± 170
		I^-	2600 ± 400

[a] Due to the nature of the experiments, the Na^+ and Cl^- values are the same data reported twice (duplicated from Table 3-1, for comparison).
[b] Supplied as the chloride salt with 330 mM of each cation.
[c] Mean and standard deviation reported are from two independent experiments.
[d] Supplied as the sodium salt at 330 mM Na^+.

Table B-2: Effects of mutation to *At*PRORP2 PPRs on affinity for pre-tRNA.

*At*PRORP1 Residue	*At*PRORP2 Variant	pre-tRNA^{Gly}		pre-tRNA^{Asp}	Highest Fold-WT
		K_D (nM)[a,b]	K_D (nM)[a,c]	K_D (nM)[a,b]	
–	WT	58 ± 8	260 ± 30	24 ± 5	1.0
A98	T31A	60 ± 8	530 ± 50		2.0
K101	L34A	28 ± 6	170 ± 10		0.7
D105	N38A	160 ± 40	1410 ± 150	70 ± 20	5.4
S131	S65A	100 ± 20	590 ± 80	35 ± 8	2.3
Y133	Q67A	70 ± 10	390 ± 70	30 ± 5	1.5
Y140	Y74S	1340 ± 170		110 ± 30	23
	Y74F	170 ± 20		45 ± 9	2.9
N175	N108A	190 ± 30	720 ± 50	70 ± 20	3.3
T180	T113A	70 ± 20	900 ± 130	30 ± 3	3.5
R210	R145A[d]	2000 ± 800			34
F119	L154A	60 ± 10	230 ± 20		1.0
E245	E180A	35 ± 4	150 ± 20		0.6
A250	S185A	40 ± 6	350 ± 40		1.3
K254	K189A	52 ± 9	350 ± 30		1.3
S280	S215A	66 ± 21	360 ± 60		1.4
S282	E217A	56 ± 14	260 ± 20		1.0
D285	K220A	129 ± 64	250 ± 40		2.2
E289	E224A	55 ± 7	310 ± 35		1.2

[a] Value and error reported are from fitting a hyperbola to the results of three independent experiments plotted together.
[b] Measured in 30 mM MOPS pH 7.8, 1 mM DTT, 150 mM NaCl, and 6 mM $CaCl_2$.
[c] Measured in 30 mM MOPS pH 7.8, 1 mM DTT, 330 mM NaCl, and 6 mM $CaCl_2$.
[d] Low protein yield for R145A limited us to one binding assay. The average and standard error from fitting a hyperbola to those data are presented.

Table B-3: Effects of mutations on *At*PRORP1 affinity for pre-tRNA.

Variant	pre-tRNA^{Cys}				pre-tRNA^{Asp}	
	K_D (nM)[a]	Fold-WT	K_D (nM)[b]	Fold-WT	K_D (nM)[a]	Fold-WT
WT	200 ± 40	1	1720 ± 160	1	160 ± 20	1
Y140A[c]	150 ± 20	0.75	>160000	>93	29600 ± 1200	170
R184A	170 ± 40	0.85	>6500	>3.8	10500 ± 2500	70

[a] Mean and standard deviation reported are from two independent experiments measured in 30 mM MOPS pH 7.8, 1 mM TCEP, 100 mM NaCl, and 1 mM $CaCl_2$.
[b] Due to high error in the experiments with Y140A and R184A under these conditions, values reported were obtained by fitting a hyperbola to the results of two independent experiments plotted together. Measured in 30 mM MOPS pH 7.8, 1 mM TCEP, 330 mM NaCl, 20 mM $CaCl_2$.
[c] We could not reach saturating enzyme conditions for Y140A in 330 mM NaCl and 20 mM $CaCl_2$. A lower limit for K_D was estimated from the available data, assuming the saturating anisotropy was the same as for WT.

Table B-4: Effects of *B*. *subtilis* pre-tRNA^{Asp} mutants on affinity for *At*PRORP1.

Variant	K_D (nM)[a]	Fold-WT
WT	160 ± 20	1
U16A	150 ± 70	0.94
U17A	110 ± 30	0.69
U20A	200 ± 20	1.25
U21A	300 ± 60	1.88

a: Mean and standard deviation reported are from two independent experiments at 330 mM NaCl.

Figures

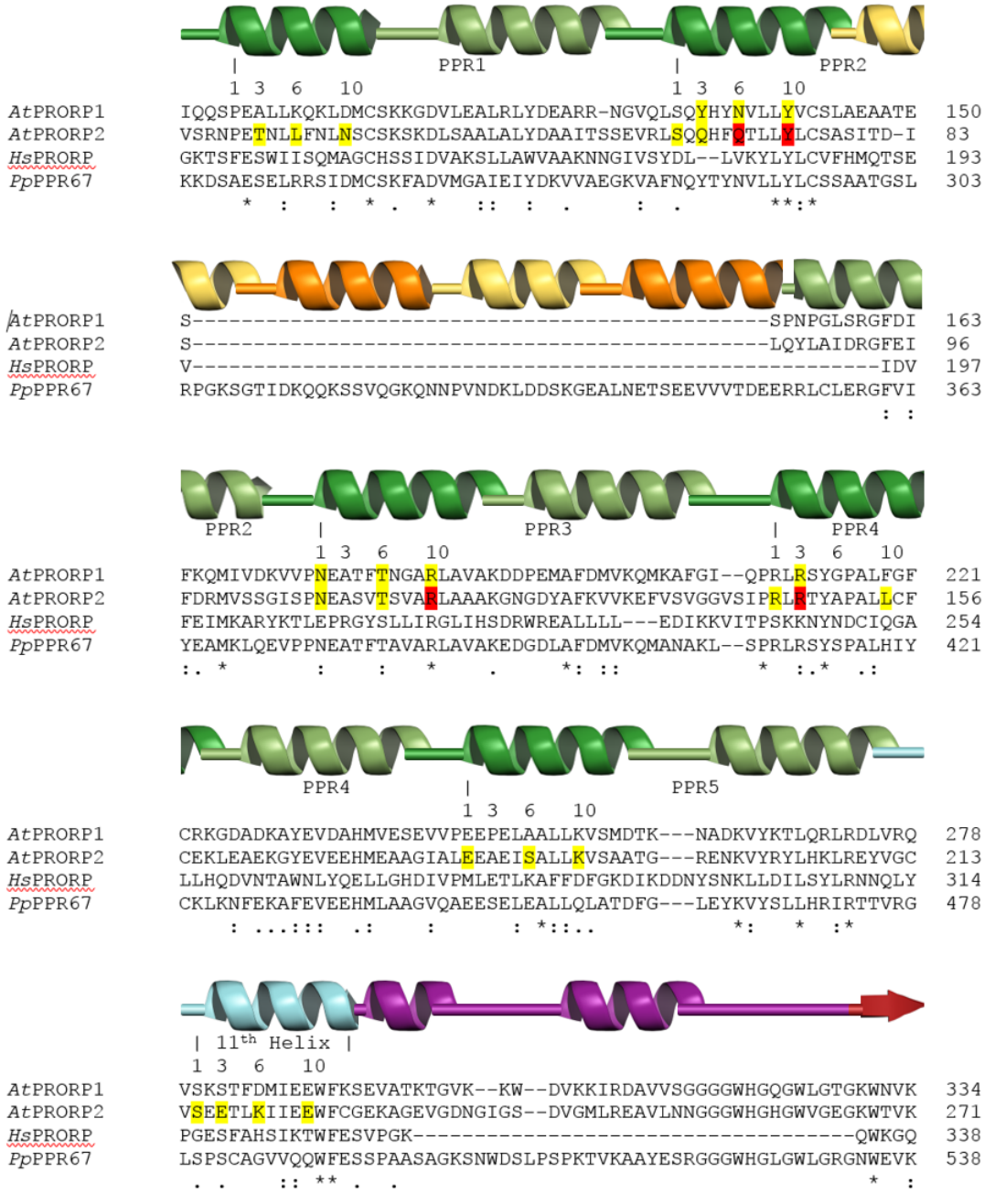


Figure B-1: Alignment of PRORP PPR domains (Clustal Omega) (60). The three Arabidopsis thaliana PRORPs (AtPRORP1–3) were aligned to the human PRORP (HsPRORP) and the Physcomitrella patens PRORPs (PpPPR63, PpPPR67, and PpPPR104). For clarity, AtPRORP3, PpPPR67, and PpPPR104 are excluded from the figure. The boundaries of the additional P. patens PRORP PPR helices were assigned using models generated with AtPRORP1 (PDB 4g23) and chloroplast PPR protein 10 (PDB 4m59) as templates in the SWISS-MODEL ExPASy server (61-63). PPR A helices in AtPRORPs and HsPRORP are colored dark green, while PPR B helices are colored pale green. Additional PPR A helices in PpPPR67 are colored orange, while PPR B helices are colored beige. PPR helix 11 is colored cyan, while the plant-specific helical insert is purple and a central domain β-strand is red. AtPRORP1 and AtPRORP2 residues mutated in this manuscript are highlighted in yellow. AtPRORP2 residues for which alanine mutants were insoluble are highlighted in red.

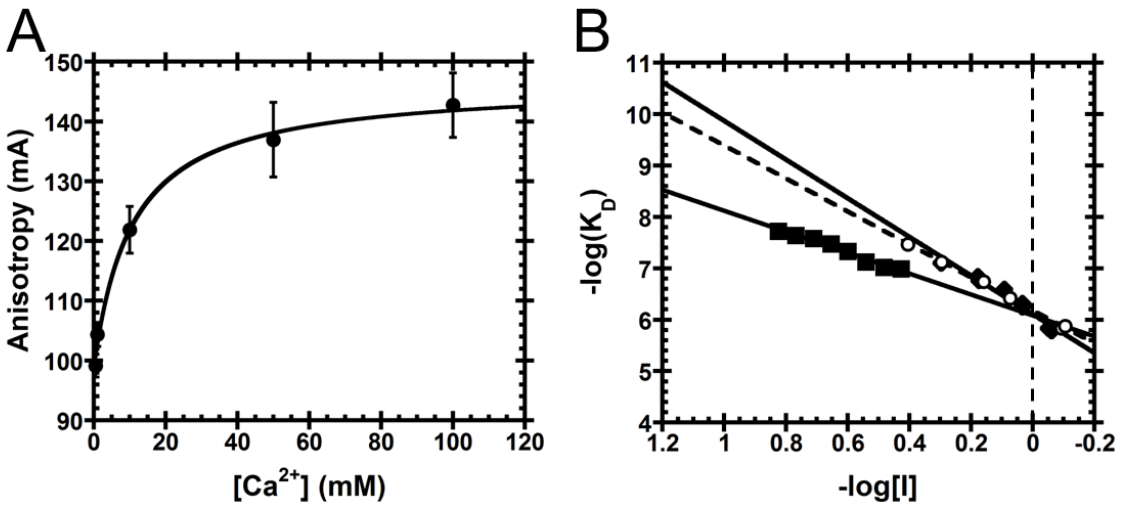


Figure B-2: Dependence of binding on anions. (**A**) Anisotropy of *Bacillus subtilis* pre-tRNA^{Asp} in the absence of PRORP is dependent on $CaCl_2$ concentration. Data reported as the mean and standard deviation of four independent experiments. A hyperbola (Equation 3-1, Materials and Methods) was fit to the data ($K_{D,app}$ = 11 ± 3). (**B**) Ionic strength (I, calculated using Equation B-3, for which *c* is the ion concentration and *z* is the ion charge) dependence of *At*PRORP1 binding to pre-tRNA^{Asp}, plotted as the –log(K_D) versus –log[I]. Equation 3-4 (Materials and Methods) was fit to the data with φ^{Na} = 0.88 or ϕ^{Ca} = 0.47 (25). Data include *At*PRORP1 binding in NaCl (♦; Z = 4.3 ± 0.3, log(K_0) = 6.11 ± 0.04), Na_2SO_4 (○; Z = 3.6 ± 0.1, log(K_0) = 6.18 ± 0.02), and $CaCl_2$ (■; Z = 4.3 ± 0.3, log(K_0) = 6.08 ± 0.09). The slope of the line ($Z\varphi$ or $Z\phi$) reports on the number of ionic interactions made to substrate phosphates proportional to fraction bound by the cation. The intercept log(K_0) (indicated by the vertical dashed line) reports on the non-ionic contributions to binding.

$$I = \frac{1}{2}\sum_{i=1}^{n}(c_i z_i^2) \qquad \text{(B-3)}$$

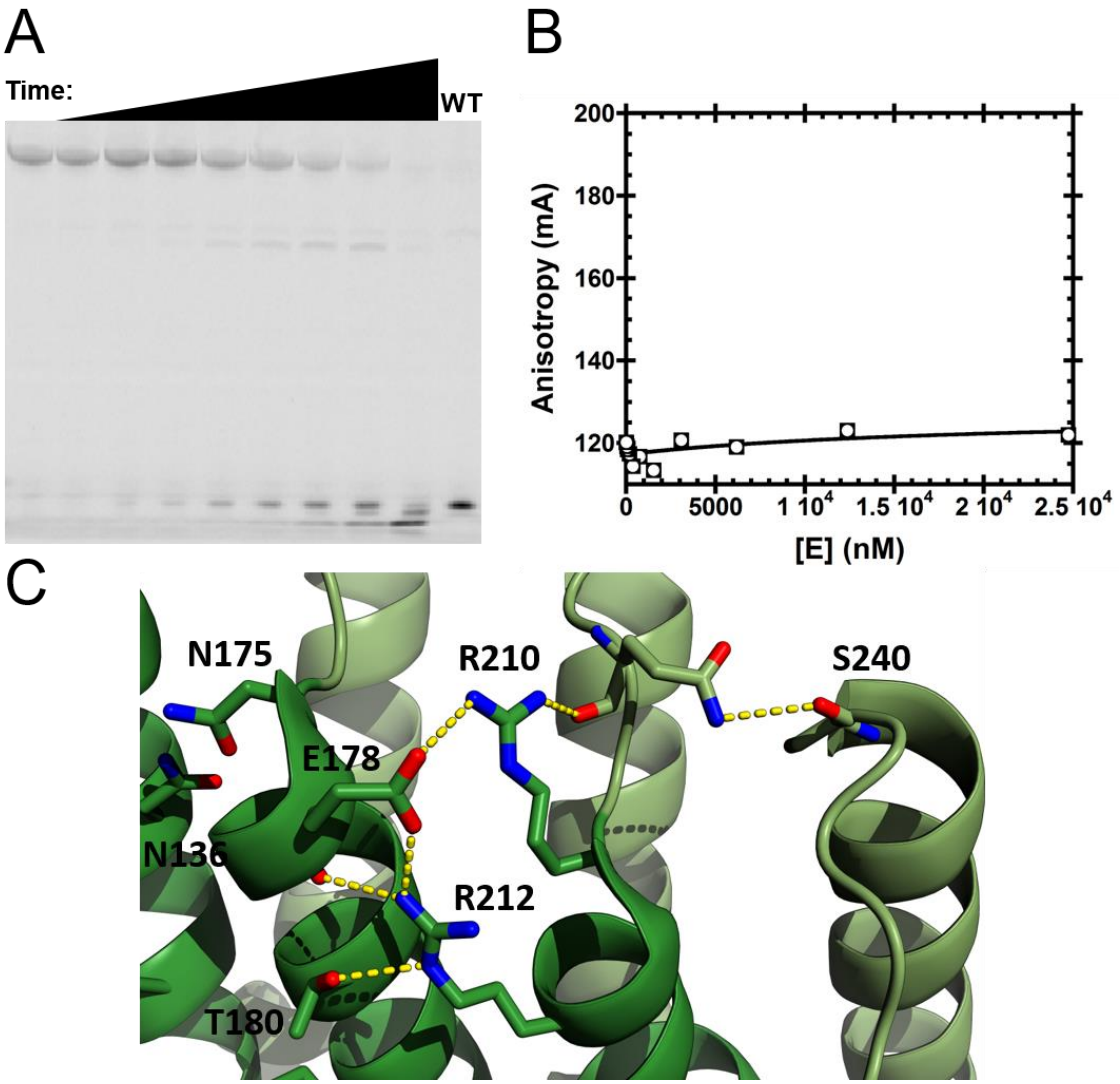


Figure B-3: Defects in R212A *At*PRORP1 catalysis and substrate binding with *B. subtilis* pre-tRNA^{Asp}. (**A**) Single turnover assay with 37 μM R212A and 30 nM substrate. Time points from 0 (left) to 21 hours were quenched and then run on high resolution urea-PAGE. 5 μM WT enzyme at 30 minutes (right) was included for comparison. (**B**) Anisotropy pre-tRNA^{Asp} binding to R212A. A hyperbola (Equation 3-1, Materials and Methods) was fit to the data ($K_{D,app}$ > 300000 nM; lower limit estimated assuming the change in anisotropy for fully-bound substrate is the same as WT). (**C**) Arg 212 in *At*PRORP1 takes part in an extended series of interactions involving both main chain and side chain atoms, beginning with Thr 180 (PPR 3 helix A) and terminating with Ser 240 (PPR 4 helix B). Generated in PyMOL (PDB 4g24) (16).

REFERENCES

1. Howard MJ, Liu X, Lim WH, Klemm BP, Fierke CA, Koutmos M, Engelke DR. RNase P enzymes: divergent scaffolds for a conserved biological reaction. *RNA Biol*. 2013;10(6):909-14.
2. Guerrier-Takada C, Gardiner K, Marsh T, Pace NR, Altman S. The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell*. 1983;35(3 Pt 2):849-57.
3. Walker SC, Engelke DR. Ribonuclease P: the evolution of an ancient RNA enzyme. *Crit Rev Biochem Mol Biol*. 2006;41(2):77-102.
4. Kurz JC, Fierke CA. The affinity of magnesium binding sites in the *Bacillus subtilis* RNase P•pre-tRNA complex is enhanced by the protein subunit. *Biochemistry*. 2002;41(30):9545-58.
5. Kurz JC, Niranjanakumari S, Fierke CA. Protein component of *Bacillus subtilis* RNase P specifically enhances the affinity for precursor-tRNA^{Asp}. *Biochemistry*. 1998;37(8):2393-400.
6. Holzmann J, Frank P, Löffler E, Bennett KL, Gerner C, Rossmanith W. RNase P without RNA: identification and functional reconstitution of the human mitochondrial tRNA processing enzyme. *Cell*. 2008;135(3):462-74.
7. Gobert A, Gutmann B, Taschner A, Gößringer M, Holzmann J, Hartmann RK, Rossmanith W, Giegé P. A single *Arabidopsis* organellar protein has RNase P activity. *Nat Struct Mol Biol*. 2010;17(6):740-4.
8. Lai LB, Bernal-Bayard P, Mohannath G, Lai SM, Gopalan V, Vioque A. A functional RNase P protein subunit of bacterial origin in some eukaryotes. *Mol Genet Genomics*. 2011;286(5-6):359-69.
9. Taschner A, Weber C, Buzet A, Hartmann RK, Hartig A, Rossmanith W. Nuclear RNase P of *Trypanosoma brucei*: a single protein in place of the multicomponent RNA-protein complex. *Cell Rep*. 2012;2(1):19-25.
10. Vilardo E, Nachbagauer C, Buzet A, Taschner A, Holzmann J, Rossmanith W. A subcomplex of human mitochondrial RNase P is a bifunctional methyltransferase–extensive moonlighting in mitochondrial tRNA biogenesis. *Nucleic Acids Res*. 2012;40(22):11583-93.
11. Gutmann B, Gobert A, Giegé P. PRORP proteins support RNase P activity in both organelles and the nucleus in *Arabidopsis*. *Genes Dev*. 2012;26(10):1022-7.
12. Howard MJ, Klemm BP, Fierke CA. Mechanistic Studies Reveal Similar Catalytic Strategies for Phosphodiester Bond Hydrolysis by Protein-only and RNA-dependent Ribonuclease P. *J Biol Chem*. 2015;290(21):13454-64.
13. Howard MJ, Lim WH, Fierke CA, Koutmos M. Mitochondrial ribonuclease P structure provides insight into the evolution of catalytic strategies for precursor-tRNA 5 ' processing. *Proc Nat Acad Sci U S A*. 2012;109(40):16149-54.
14. Imai T, Nakamura T, Maeda T, Nakayama K, Gao X, Nakashima T, Kakuta Y, Kimura M. Pentatricopeptide repeat motifs in the processing enzyme PRORP1 in *Arabidopsis thaliana* play a crucial role in recognition of nucleotide bases at TψC loop in precursor tRNAs. *Biochem Biophys Res Commun*. 2014;450(4):1541-6.

15. Anantharaman V, Aravind L. The NYN Domains: Novel Predicted RNAses with a PIN Domain-like Fold. *RNA Biol*. 2006;3(1):18-27.
16. Schrödinger, LLC. PyMOL Molecular Graphics System, Version 1.52010.
17. Pavlova LV, Gößringer M, Weber C, Buzet A, Rossmanith W, Hartmann RK. tRNA processing by protein-only versus RNA-based RNase P: kinetic analysis reveals mechanistic differences. *ChemBioChem*. 2012;13(15):2270-6.
18. Gobert A, Pinker F, Fuchsbauer O, Gutmann B, Boutin R, Roblin P, Sauter C, Giegé P. Structural insights into protein-only RNase P complexed with tRNA. *Nat Commun*. 2013;4(1353):1-8.
19. Brillante N, Gößringer M, Lindenhofer D, Toth U, Rossmanith W, Hartmann RK. Substrate recognition and cleavage-site selection by a single-subunit protein-only RNase P. *Nucleic Acids Res*. 2016;44(5):2323-36.
20. Howard MJ, Karasik A, Klemm BP, Mei C, Shanmuganathan A, Fierke CA, Koutmos M. Differential substrate recognition by isozymes of plant protein-only Ribonuclease P. *RNA*. 2016;22(5):782-92.
21. Barkan A, Rojas M, Fujii S, Yap A, Chong YS, Bond CS, Small I. A Combinatorial Amino Acid Code for RNA Recognition by Pentatricopeptide Repeat Proteins. *PLOS Genetics*. 2012;8(8):e1002910.
22. Yagi Y, Hayashi S, Kobayashi K, Hirayama T, Nakamura T. Elucidation of the RNA recognition code for pentatricopeptide repeat proteins involved in organelle RNA editing in plants. *PLOS One*. 2013;8(3):e57286.
23. Record MT, Jr., Anderson CF, Lohman TM. Thermodynamic analysis of ion effects on the binding and conformational equilibria of proteins and nucleic acids: the roles of ion association or release, screening, and ion effects on water activity. Q Rev Biophys. 1978;11(2):103-78.
24. Barkley MD, Lewis PA, Sullivan GE. Ion effects on the lac repressor--operator equilibrium. *Biochemistry*. 1981;20:3842-51.
25. deHaseth PL, Lohman TM, Record MT, Jr. Nonspecific interaction of lac repressor with DNA: an association reaction driven by counterion release. *Biochemistry*. 1977;16(22):4783-90.
26. Karasik A, Shanmuganathan A, Howard MJ, Fierke CA, Koutmos M. Nuclear Protein-Only Ribonuclease P2 Structure and Biochemical Characterization Provide Insight into the Conserved Properties of tRNA 5' End Processing Enzymes. *J Mol Biol*. 2016;428(1):26-40.
27. Day-Storms JJ, Niranjanakumari S, Fierke CA. Ionic interactions between PRNA and P protein in *Bacillus subtilis* RNase P characterized using a magnetocapture-based assay. *RNA*. 2004;10(10):1595-608.
28. Record MT, Jr., Lohman ML, De Haseth P. Ion effects on ligand-nucleic acid interactions. *J Mol Biol*. 1976;107(2):145-58.
29. Guckian KM, Schweitzer BA, Ren RX, Sheils CJ, Tahmassebi DC, Kool ET. Factors Contributing to Aromatic Stacking in Water: Evaluation in the Context of DNA. *J Am Chem Soc*. 2000;122:2213-22.
30. Leroy JL, Guéron M, Thomas G, Favre A. Role of divalent ions in folding of tRNA. *Eur J Biochem*. 1977;74:567-74.

31. Liu X, Chen Y, Fierke CA. A real-time fluorescence polarization activity assay to screen for inhibitors of bacterial ribonuclease P. *Nucleic Acids Res*. 2014;42(20):e159.
32. Anderson P, Bauer W. Supercoiling in closed circular DNA: dependence upon ion type and concentration. *Biochemistry*. 1978;17:594-601.
33. Gebala M, Giambasu GM, Lipfert J, Bisaria N, Bonilla S, Li G, York DM, Herschlag D. Cation-Anion Interactions within the Nucleic Acid Ion Atmosphere Revealed by Ion Counting. *J Am Chem Soc*. 2015;137(46):14705-15.
34. Crary SM, Niranjanakumari S, Fierke CA. The protein component of *Bacillus subtilis* ribonuclease P increases catalytic efficiency by enhancing interactions with the 5' leader sequence of pre-tRNA^{Asp}. *Biochemistry*. 1998;37(26):9409-16.
35. Rueda D, Hsieh J, Day-Storms JJ, Fierke CA, Walter NG. The 5' leader of precursor tRNA^{Asp} bound to the *Bacillus subtilis* RNase P holoenzyme has an extended conformation. *Biochemistry*. 2005;44(49):16130-9.
36. Kobayashi K, Kawabata M, Hisano K, Kazama T, Matsuoka K, Sugita M, Nakamura T. Identification and characterization of the RNA binding surface of the pentatricopeptide repeat protein. *Nucleic Acids Res*. 2012;40(6):2712-23.
37. Yin P, Li Q, Yan C, Liu Y, Liu J, Yu F, Wang Z, Long J, He J, Wang H-W, Wang J, Zhu J-K, Shi Y, Yan N. Structural basis for the modular recognition of single-stranded RNA by PPR proteins. *Nature*. 2013;504(7478):168-71.
38. Small ID, Peeterns N. The PPR motif – a TPR-related motif prevalent in plant organellar proteins. *Trends Biochem Sci*. 2000;25(2):46-7.
39. Schmitz-Linneweber C, Small I. Pentatricopeptide repeat proteins: a socket set for organelle gene expression. *Trends Plant Sci*. 2008;13:663-70.
40. Ringel R, Sologub M, Morozov YI, Litonin D, Cramer P, Temiakov D. Structure of human mitochondrial RNA polymerase. *Nature*. 2011;478:269-73.
41. O'Toole N, Hattori M, Andres C, Iida K, Lurin C, Schmitz-Linneweber C, Sugita M, Small I. On the expansion of the pentatricopeptide repeat gene family in plants. *Mol Biol Evol*. 2008;25(6):1120-8.
42. Chen X, Feng F, Qi W, Xu L, Yao D, Wang Q, Song R. Dek35 encodes a PPR protein that affects cis-splicing of mitochondrial nad4 intron 1 and seed development in maize. *Mol Plant*. 2016;S1674-2052(16):30189-7.
43. Reinhard L, Sridhara S, Hallberg BM. Structure of the nuclease subunit of human mitochondrial RNase P. *Nucleic Acids Res*. 2015;43(11):5664-72.
44. Jones S, Daley DT, Luscombe NM, Berman HM, Thornton JM. Protein-RNA interactions: a structural analysis. *Nucleic Acids Res*. 2001;29:943-54.
45. Singh M, Wang Z, Koo BK, Patel A, Cascio D, Collins K, Feigon J. Structural basis for telomerase RNA recognition and RNP assembly by the holoenzyme La family protein p65. *Mol Cell*. 2012;47:16-26.
46. Reiter NJ, Osterman A, Torres-Larios A, Swinger KK, Pan T, Mondragón A. Structure of a bacterial ribonuclease P holoenzyme in complex with tRNA. *Nature*. 2010;468(7325):784-9.
47. Sugita C, Komura Y, Tanaka K, Kometani K, Satoh H, Sugita M. Molecular characterization of three PRORP proteins in the moss *Physcomitrella patens*: nuclear PRORP protein is not essential for moss viability. *PLOS One*. 2014;9(10):e108962.

48. Sugita M, Ichinose M, Ide M, Sugita C. Architecture of the PPR gene family in the moss *Physcomitrella patens*. *RNA Biol*. 2013;10(9):1439-45.
49. Shi H, Moore PB. The crystal structure of yeast phenylalanine tRNA at 1.93 A resolution: a classic structure revisited. *RNA*. 2000;6(8):1091-105.
50. Hoang C, Ferré-D'Amaré AR. Cocrystal structure of a tRNA ψ55 pseudouridine synthase: nucleotide flipping by an RNA-modifying enzyme. *Cell*. 2002;107:929-39.
51. Zhang J, Ferré-D'Amaré AR. The tRNA Elbow in Structure, Recognition and Evolution. *Life*. 2016;6(1)3.
52. Baird TT, Jr., Waheed A, Okuyama T, Sly WS, Fierke CA. Catalysis and inhibition of human carbonic anhydrase IV. *Biochemistry*. 1997;36(9):2669-78.
53. Barkley MD. Salt dependence of the kinetics of the lac repressor-operator interaction: role of nonoperator deoxyribonucleic acid in the association reaction. *Biochemistry*. 1981;20(13):3833-42.
54. Hsieh J, Fierke CA. Conformational change in the *Bacillus subtilis* RNase P holoenzyme–pre-tRNA complex enhances substrate affinity and limits cleavage rate. *RNA*. 2009;15(8):1565-77.
55. Hsieh J, Koutmou KS, Rueda D, Koutmos M, Walter NG, Fierke CA. A divalent cation stabilizes the active conformation of the *B*. *subtilis* RNase P•pre-tRNA complex: a role for an inner-sphere metal ion in RNase P. *J Mol Biol*. 2010;400(1):38-51.
56. Hutchison CA, Phillips S, Edgell MH, Gillam S, Jahnke P, Smith M. Mutagenesis at a specific position in a DNA sequence. *J Biol Chem*. 1978;253(18):6551-60.
57. Milligan JF, Uhlenbeck OC. Synthesis of small RNAs using T7 RNA polymerase. *Methods Enzymol*. 1989;180:51-62.
58. Pierce BG, Wiehe K, Hwang H, Kim BH, Vreven T, Weng Z. ZDOCK server: interactive docking prediction of protein-protein complexes and symmetric multimers. *Bioinformatics*. 2014;30(12):1771-3.
59. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. *Acta Crystallographica D Biol Crystallogr*. 2010;66(Pt 4):486-501.
60. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Soding J, Thompson JD, Higgins DG. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol*. 2011;7:539.
61. Arnold K, Bordoli L, Kopp J, Schwede T. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics*. 2006;22(2):195-201.
62. Kiefer F, Arnold K, Kunzli M, Bordoli L, Schwede T. The SWISS-MODEL Repository and associated resources. *Nucleic Acids Res*. 2009;37(Database issue):D387-92.
63. Guex N, Peitsch MC, Schwede T. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: a historical perspective. *Electrophoresis*. 2009;30 Suppl 1:S162-73.

CHAPTER 4

Developing cross-linking methods for mapping PRORP-substrate complexes

ABSTRACT

Ribonuclease P (RNase P) enzymes are endonucleases responsible for the 5' end maturation of precursor transfer ribonucleic acids (pre-tRNAs). RNase P exists as an RNA-based enzyme in all domains of life. In addition, protein-only RNase P (PRORP) is found in many eukaryotes and localizes to various cellular compartments. Here, we develop a method using non-natural amino acid (NNAA) cross-linkers to identify sites of contact between an *Arabidopsis thaliana* PRORP and its pre-tRNA substrates. The method makes use of an Amber codon suppression system to install photo-activatable cross-linkers in the protein, allowing for the formation of covalent complexes. When we irradiate *E*. *coli* expressing PRORP-containing photo-cross-linkers, we observe an increase in the apparent molecular weight for NNAA-substituted PRORPs by denaturing PAGE that is consistent with a PRORP-tRNA complex. Furthermore, the NNAA-substituted proteins are stable and generate a specific cross-linked species with pre-tRNA *in vitro* upon irradiation. We use a primer extension assay to identify sites on the pre-tRNA that cross-link to the protein. Our goal in developing this method is to map the PRORP-tRNA interaction surface for the PRORPs from plants and humans, but it could be easily adapted for use with other RNA-binding proteins.

INTRODUCTION

Ribonuclease P (RNase P) is a metal-dependent endonuclease required for the 5' end maturation of precursor tRNAs (pre-tRNAs) (1). RNase P is classically known as an RNA-dependent macromolecular catalyst, one of only a few capable of completing multiple turnovers (2). Many eukaryotes also contain protein-only RNase Ps (PRORPs) that are required for 5' end maturation activity in various tRNA-coding cellular compartments (3). Humans require an RNA-dependent RNase P for nuclear pre-tRNA maturation and a three-component PRORP for mitochondrial pre-tRNA maturation (4).

In contrast, the homologous PRORPs from land plants are single-subunit enzymes that catalyze maturation of pre-tRNAs from all compartments without additional protein or RNA cofactors (5, 6). PRORPs utilize a two-metal ion catalytic strategy that is similar to that of the RNA-dependent RNase P (7). Many important details about PRORP-substrate recognition are already known. The D-loop (8-10), as well as the lengths of the TψC-arm and acceptor arm (9), are structural features required for substrate recognition. PRORPs contain a pentatricopeptide repeat (PPR) domain, a broad class of domains found in RNA-binding proteins, that are responsible for a significant portion of the substrate affinity (11, 12). In Chapter 3 we identified the surface of a PRORP substrate-recognition domain that is responsible for recognizing substrate. While complementary mutation experiments have been attempted (9), none have successfully identified the sites of direct contact between PRORPs and their substrates.

In this chapter, we develop a method to identify the sites of contact between PRORPs and pre-tRNA substrates. We make use of the Amber suppression system developed by Peter Schulz to site-specifically incorporate non-natural amino acid (NNAA) cross-linkers into *Arabidopsis thaliana* PRORP1 (13, 14). We previously identified multiple tyrosine residues positioned on the pre-tRNA binding surface that are important for substrate binding (Chapter 3). To make substitutions at these sites, we selected two NNAAs of similar structure, *para*-azido-L-phenylalanine (pAzF) and *para*-benzoyl-L-phenylalanine (pBF; Fig. 4-1), to

minimize the impact of the mutations on binding affinity. The pBF substitution has previously been used to form cross-links with nucleic acid-binding proteins (15).

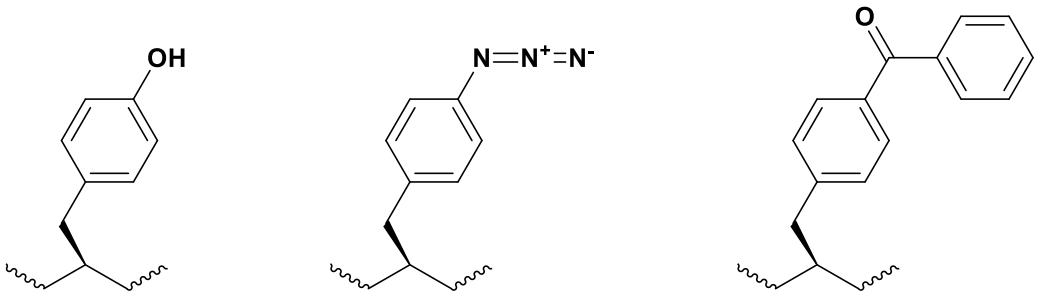


Figure 4-1: NNAA structures. (**Left**) tyrosine, for comparison. (**Middle**) *para*-azido-L-phenylalanine (pAzF). (**Right**) *para*-benzoyl-L-phenylalanine (pBF).

Our method uses primer extension in a reverse transcription (RT) assay to identify cross-linked sites between protein-only RNase Ps and pre-tRNA substrates (Fig. 4-2). This method could be applied to other RNA-binding proteins, including the human mitochondrial PRORP. This would be an interesting application because the three-component system is more complicated and few details about the complex architecture are currently known. It should also be possible to generate protein-protein cross-links to identify protein-protein interaction surfaces within multi-protein complexes, such as the human PRORP. This method could also be used to test the predictive model we presented in Chapter 3.

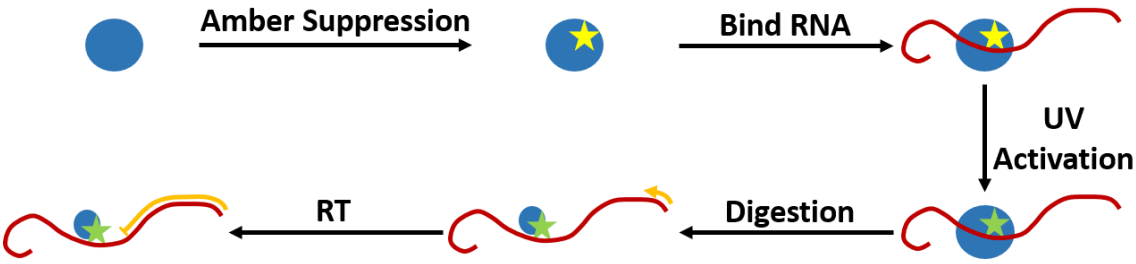


Figure 4-2: Work-flow for cross-linking method. NNAAs (star) are incorporated into the protein of interest (blue circle) by the Amber suppression system. Protein is bound to RNA (red line), followed by photo-activated cross-linking with UV, protein digestion with protease, and cross-linking site detection by reverse transcription.

RESULTS

Expressing non-natural amino acid-substituted *At*PRORP1

Expression lines (BL21/pETM-11/pEVOL) were prepared as described in the methods. We performed small-scale expression tests to determine the requirement of each expression component and overall efficiency (Fig. 4-3). Three components of the induction media should be required to obtain full length Δ76PRORP1 where the gene contains an Amber mutation: isopropyl β-D-1-thiogalactopyranoside (IPTG), arabinose, and either pAzF or pBF. IPTG induces T7 RNA polymerase which transcribes the PRORP1 gene, arabinose induces expression for the mutant aminoacyl-tRNA synthetase (muRS), and pAzF or pBF are needed for the muRS to charge the Amber suppressor tRNA ($tRNA^{SUP}$) to obtain efficient read-through of the Amber stop codon.

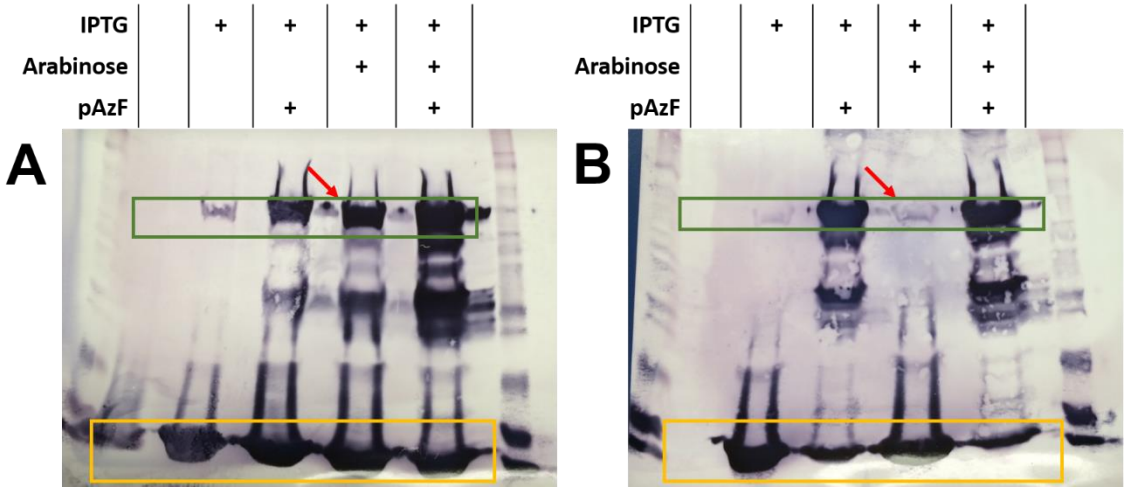


Figure 4-3: α-His Western blots from expression testing. Green boxes indicate full-length PRORP1. Orange boxes indicate the protein with failed Amber read-through, resulting in an N-terminal truncation stopped at the Amber codon. Red arrows highlight the requirement for pAzF to obtain Amber suppression. Ladders flank the experimental lanes in both gels. (**A**) Expression results in LB medium. pAzF was not required for Amber suppression. (**B**) Expression results in LRMM. pAzF was required for Amber suppression.

We noted that when expressing in Luria-Bertani (LB) complete medium, we obtained full-length PRORP1 even when the induction media lacked pAzF (Fig. 4-3A, red arrow). This indicates that we obtained efficient read-through of the Amber codon, even without $tRNA^{SUP}$ charged with pAzF. We hypothesized that the muRS, which was based on a tyrosine RS, may be able to utilize the tyrosine, which is abundant in LB. Thus, we limited tyrosine by culturing cells in the LeMaster and

Richard's minimal media (LRMM, Fig. 4-3B) (16), which forces *E. coli* to synthesize their own amino acids and thus should limit the intracellular tyrosine concentration. Interestingly, we obtained read-through when pAzF was supplied, regardless of whether we included arabinose in the induction media. We hypothesize that an *E. coli* tRNA synthetase can charge the $tRNA^{SUP}$ (possibly a TyrRS) with pAzF. In support of this hypothesis, both arabinose and pBF were strictly required for Amber suppression using the pBF muRS (data not shown). Regardless, we observe a slight increase in read-through for pAzF with the addition of arabinose. Given that we observe no other defects in growth or expression, we include arabinose in all further induction media.

Because *At*PRORP1 can bind to bacterial pre-tRNAs *in vitro* (10), we tested the ability of the NNAA-substituted PRORP1 to cross-link in the BL21(DE3) *E. coli* expression cells. We grew small-scale expressions in parallel, induced His_6-PRORP expression in the presence of either pAzF or pBF overnight, irradiated the cells, then lysed and ran samples from each on a Western blot, probing with an anti-His_6 antibody (Fig. 4-4). Significant amounts of cross-linked species resulted from the Y133 and Y140 mutants for both pAzF and pBF, while the T180 mutants and WT PRORP1 (Fig. C-1) did not.

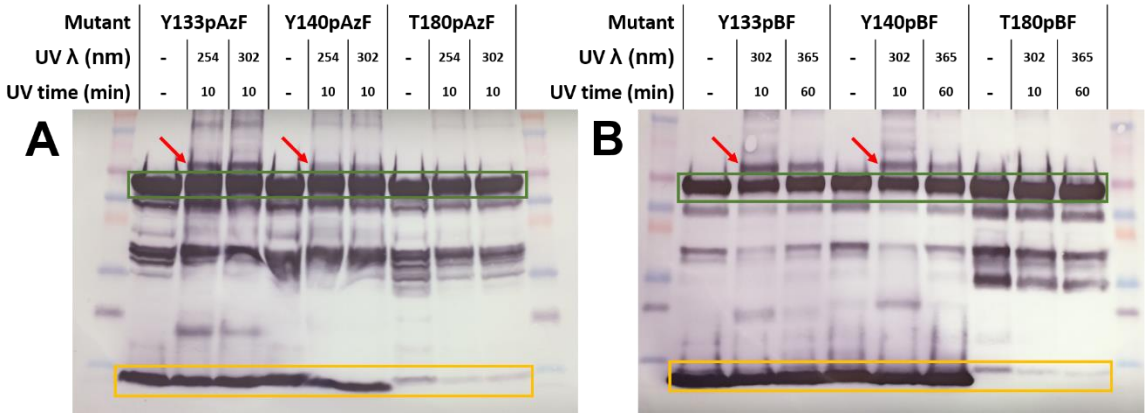


Figure 4-4: α-His Western blots from cross-linking experiments in *E. coli*. Green boxes indicate full-length PRORP1. Orange boxes indicate the protein resulting from failed Amber read-through. Red arrows indicate the MW of putative PRORP1-tRNA cross-linked species (≈75kDa). Ladders flank the experimental lanes in both gels. (**A**) PRORP1 with Amber mutations at Y133, Y140, and T180 expressed in the presence of pAzF. The aryl-azide was activated by irradiation at either 254 or 302 nm for 10 minutes (λ_{max} < 310). (**B**) PRORP1 with Amber mutations at Y133, Y140, and T180 expressed in the present of pBF. The benzophenone was irradiated with UV at either 302 nm for 10 minutes or 365 nm for 60 minutes (λ_{max} > 320).

We initially attempted to purify Y140pAzF for *in vitro* cross-linking by expressing in *E*. *coli* as we described previously for WT PRORP1 (7, 10, 11). However, the protein did not efficiently cross-link with purified pre-tRNA *in vitro* and the mass was not correct for pAzF incorporation by quantitative time-of-flight mass spectrometry (Q-TOF MS) (see the pAzF-*At*PRORP1 purification section, pg 114). We proposed that the aryl-azide was unstable in the lysis or purification buffers, likely due to reaction with the TCEP reducing agent in the buffer and moved our focus to the pBF-containing mutants.

Cross-linking pBF-substituted *At*PRORP1 *in vitro*

We verified the mass of the Y133pBF mutant produced in a test expression by Q-TOF (Fig. C-2), which showed the expected mass-shift (59541.4 Da versus 59450.8 Da). We then expressed the Y133pBF mutant at a 3 L scale and purified the protein using the WT PRORP1 protocol (7, 10, 11). The Y133pBF mutant binds pre-tRNA weaker than WT (Fig. C-3), but the affinity is tight enough to yield titration conditions in our normal cross-linking buffer, suggesting that we may be able to generate an *in vitro* protein-RNA cross-linked species.

We ran an irradiation time-course to establish optimal conditions for cross-linking (Fig. 4-5A). Samples were irradiated at 365 nm for various times, then run on urea-PAGE. For future iterations of the protocol, we should run an irradiation time course on the substrate alone, followed by reverse transcription, in order to increase the signal to noise of primer extension by establishing the point at which UV damage significantly inhibits RT. Quantification indicated inefficient cross-linking by the pBF-substituted protein, with very long times required to reach a maximum ~ 10%. We saturated substrate with Y133pBF PRORP1 at higher concentrations to obtain enough sample for the reverse transcription (RT) assays. Under these conditions, we obtained ~ 9% cross-linked species after irradiating for 2 hours (Fig. 4-5B). In the future, this experiment should include an irradiated/no enzyme control lane.

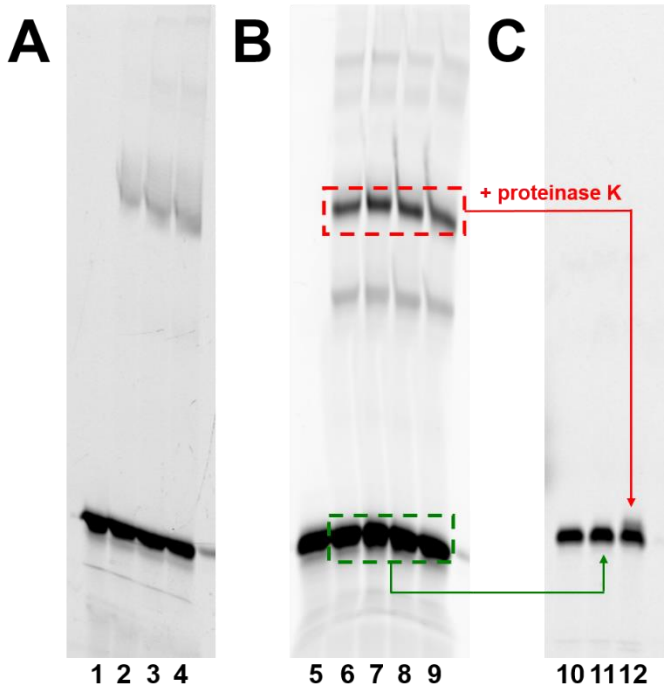


Figure 4-5: Cross-linking Y133pBF PRORP1 to 5'-fluorescein-labeled *B. subtilis* pre-tRNA^{Asp}. Fractionated by urea-SDS-PAGE gel and substrate visualized using fluorescence. Regions marked by dashed boxes were cut from the gel and re-purified. (**A**) Cross-linking time course. 0.4 μM pre-tRNA pre-incubated with 2 μM Y133pBF PRORP1 for 20 minutes. Lane 1: no UV-irradiation. Lane 2-4: UV-irradiation at 365 nm for 40 min, 76 min, and 145 min, respectively. (**B**) Cross-linking reaction. 10 μM pre-tRNA was pre-incubated with 15 μM Y133pBF PRORP1 for 20 minutes. Lane 5: no-UV control. Lanes 6-9: UV-irradiated at 365 nm for 120 minutes. (**C**) Re-purified substrates. Lane 10: substrate. Lane 11: UV-irradiated substrate (not cross-linked) re-purified from panel B. Lane 12: cross-linked substrate from panel B, re-purified and treated with proteinase K.

We purified each sample from the gel for use as the template in RT reactions. UV-irradiated substrate was re-purified from the gel as previously described (7), including eluting into Tris pH 8, EDTA, and SDS buffer (TE-SDS) overnight. Cross-linked sample was likewise re-purified, with the addition of a proteinase K treatment step prior to ethanol precipitation (proteinase K cleaves the C-terminal side of aliphatic and aromatic amino acids). To verify the integrity of the samples, we ran a comparison gel (Fig. 4-5C). The non-irradiated substrate (lane 10) and irradiated substrate (lane 11) have identical mobility, with very little degradation apparent. Furthermore, the cross-linked/proteolyzed substrate band (lane 12) was smeared, possibly due to the presence of a PRORP peptide fragment(s) (SQ[pBF]HY, assuming proteinase K cannot recognize the pBF/RNA site). In future cross-linking experiments, we should set aside a portion of the cross-linked sample to include a non-proteolyzed control lane.

Cross-linking site identification

To identify the site of cross-linking on pre-tRNA, we employed a reverse transcriptase (RTase) primer extension assay. The assay should detect sites of cross-linking by decreasing the read-through by the RTase when it encounters a protein cross-linked nucleotide. A similar assay was previously used to monitor tRNA methylation (17). We initially designed the oligo-DNA primer against the 3' end of the pre-tRNA (Fig. 4-6, primer 1). As might be expected, the assays indicate decreased read-through for the UV-irradiated substrate relative to the non-irradiated substrate (compare Fig. 4-7, lanes 2 and 3). Further, several sites have either significantly reduced or significantly increased read-through in the cross-linked, proteolyzed substrate (compare Fig. 4-7, lanes 3 and 4). The increased read-through may indicate that PRORP1 binding distorted that region of the tRNA structure, reducing the base damage that accumulated at the site. The primary location of decreased read-through is located in what we approximate to be the D-loop. However, the resolution in that region is low due to the longer extension (39–48-nt) required to reach it. Therefore, we designed a second primer to target the D-loop by pairing with the anti-codon stem (Fig. 4-6, primer 2).

Figure 4-6: Reverse transcription primer design. The *B. subtilis* pre-tRNA^{Asp} with a 5-nt leader sequence is shown. The sites of cross-linking identified in Fig. 4-7 are indicated. Key tRNA loops are indicated. Primers include a 5' fluorescein (Fl) label and are shown base-paired with the tRNA. Potential primer extension lengths emphasized by dashed red arrows. Primer 1 (used in Fig. 4-7A) has optimal resolution in the TψC-loop and anti-codon loop. Primer 2 (used in Fig. 4-7B) has optimal resolution that extends to the 5' end of the pre-tRNA.

The second primer produces better resolution in the D-loop as a result of the shorter extension (5–14-nt) required to reach it. The primary sites of read-through decrease upon UV cross-linking are at +6, which corresponds to the first nucleotide in the D-loop, and +20/+22, which correspond to U_8 in the loop between the acceptor and D-stems and G_6 in the acceptor stem (Figs. 4-6, 4-7B). Interestingly, two nucleotides in the D-stem and one in the 5' leader have increased read-through. This may indicate that PRORP1 binding distorts the pre-tRNA structure,

resulting in less UV damage to those bases. PRORP1 binding to the D-loop is consistent with previous data indicating that the D-loop is important for PRORP recognition (9, 10).

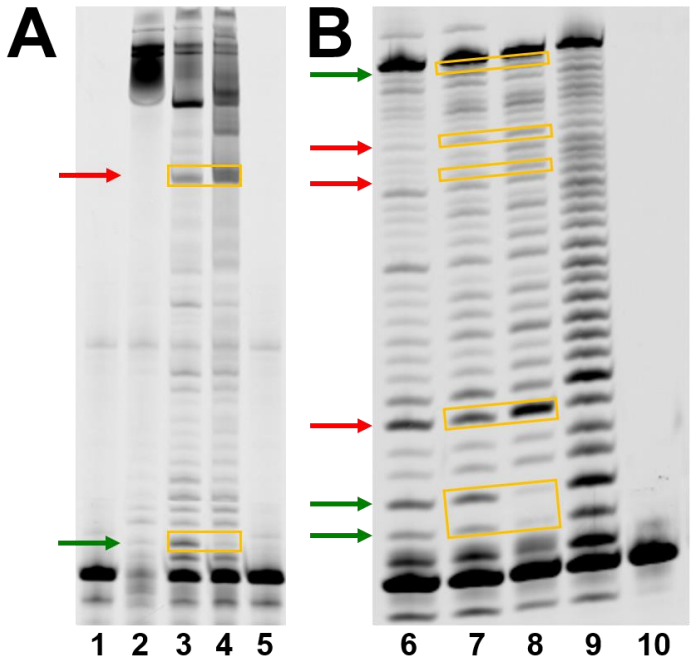


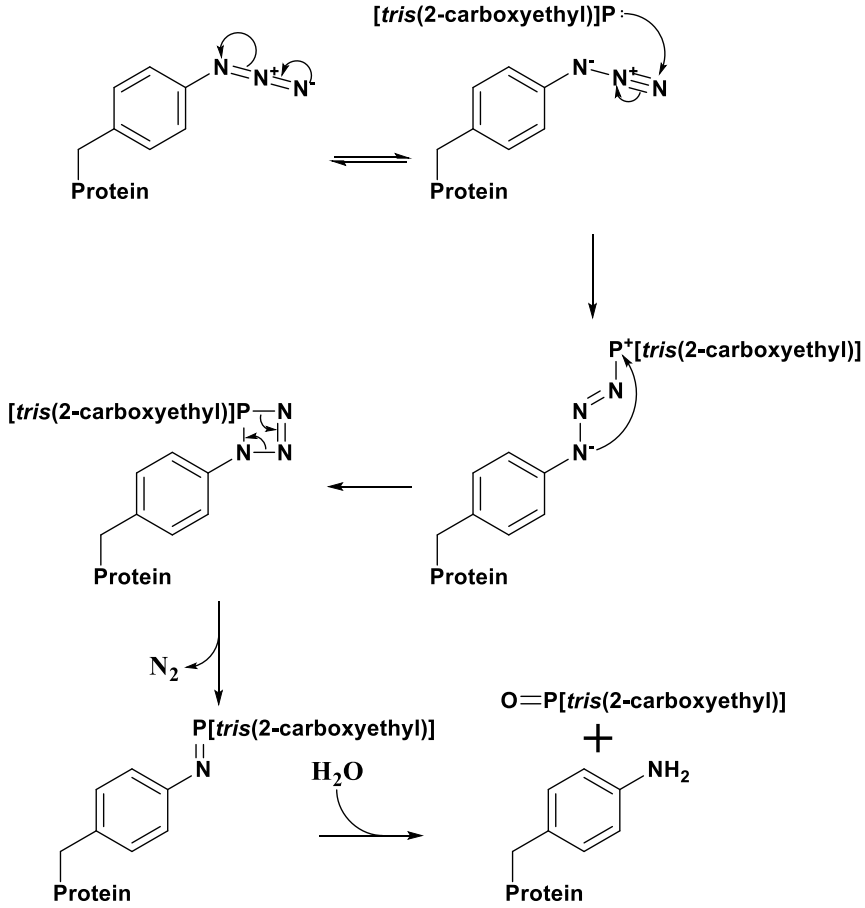
Figure 4-7: RT primer extension assays. Bands were quantified using ImageQuant and normalized to total lane intensity. Red arrows indicate the bands that increased to > 2-fold upon cross-linking with protein over UV-irradiated substrate alone. Green arrows indicate bands that decreased to < 0.5-fold in the protein cross-linked species. (**A**) Primer extension using primer 1 (Fig. 4-6). Lanes 1 & 5: primer alone. Lane 2: substrate-alone as template. Lane 3: UV-irradiated substrate-alone as template. Lane 4: cross-linked, proteolyzed substrate as template. (**B**) Primer extension using primer 2 (Fig. 4-6). Lane 6: substrate-alone as template. Lane 7: UV-irradiated substrate-alone as template. Lane 8: cross-linked, proteolyzed substrate as template. Lane 9: substrate was laddered with 10mM hydroxide in 1mM EDTA for 10 minutes, then used to template the RT reaction. Lane 10: primer alone.

pAzF-*At*PRORP1 purification

The initial purification of Y133pAzF PRORP1 yielded a protein with the incorrect mass. We purified the protein using the WT PRORP1 protocol and examined the mutant by Q-TOF MS, using WT PRORP1 as a standard (Fig. C-4). The mass detected was -1 Da from WT, while the expected mass shift is +25 Da. Our normal purification/storage buffers contain tris(2-carboxyethyl)phosphine (TCEP), which can react with the azide by the Staudinger reaction to yield a fully-reduced amine, oxidized phosphine, and N_2 (Scheme 4-1)[1]. The resulting protein

[1] Thanks to Matthew Henley for bringing this to my attention and solving a couple months' of issues.

contains a *para*-amino-L-phenylalanine (pAF), with an expected mass shift of -1 Da from WT.



Scheme 4-1: Staudinger reaction between TCEP and the aryl azide in pAzF. Products include N_2, phosphine oxide, and *para*-amino-L-phenylalanine (pAF).

Like TCEP, other common biochemical reducing agents including dithiothreitol (DTT) and β-mercaptoethanol (βME) also reduce aryl azides to an aryl amine (18) and therefore cannot be used during purification. These reactions proceed with kinetics that are sub-optimal below pH 8, with DTT reacting much faster ($t_{1/2}$ ≈ minutes) than βME (< 20% of the azide is reduced at 24 hours) (18). To test the stability of the protein in the various conditions, we lysed Y133pAzF-expressing cells in our standard buffer with differing reducing conditions (Fig. 4-8). The solubility of the protein was not affected, whether we included reducing agents or not (Fig. 4-8A). Likewise, the protein was relatively stable without a reducing agent

at 4°C for seven days (Fig. 4-8B). Thus, the best option for preparing pAzF-substituted protein in the future is to exclude reducing agents from the buffers.

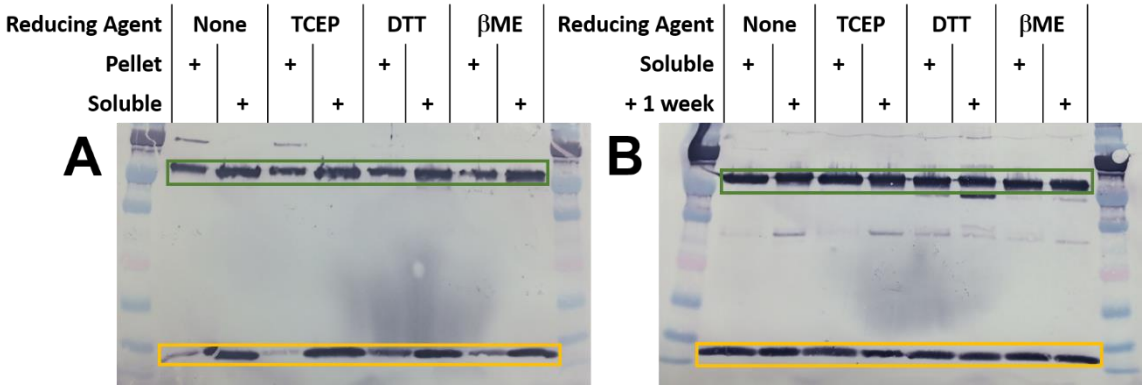


Figure 4-8: α-His Western blots from *E. coli* lysates treated with varying reducing conditions. Lysis buffers contained 1 mM of each reducing agent. Green boxes indicate full-length PRORP1, while orange boxes indicate the protein resulting from failed Amber read-through. Ladders flank the experimental lanes in both gels. (**A**) Lysis under various reducing conditions. The reducing conditions did not affect the fraction of soluble protein. (**B**) Soluble fractions from panel A (stored in loading dye at -20°C) run against the cleared lysate left at 4°C for seven days. Excluding reducing agents did not decrease protein stability.

DISCUSSION

In this chapter, we developed a method to identify sites of contact between PRORP1 and pre-tRNA substrates. This novel method was developed to obtain good resolution and specificity with the PRORP-substrate complex. We considered alternative cross-linking methods, such as incorporating the photo-activatable cross-linker 5-bromouridine (BrU) into substrates, foot-printing the PRORP-substrate complex with Fe-EDTA (19), and azidophenyacyl bromide thiol labeling (20), but each has additional caveats or complications. *At*PRORPs contain > 10 cysteines, making site-specific labeling by azidophenyacyl bromide obtained with the bacterial RNase P protein difficult and improbable (20). BrU incorporation can be obtained by replacing UTP with BrUTP in the *in vitro* transcription reaction, but the cross-linking species produced do not yield specific information about the portion of protein that was cross-linked to a given site on the RNA. Likewise, we could map the PRORP-substrate complex using Fe-EDTA foot-printing (19), but the resolution achieved would be lower than we obtained the other methods discussed and the results are likely to simply confirm the nuclease foot-printing results produced by Gobert, *et al* (8).

We obtained specific incorporation of pAzF and pBF into PRORP in a minimal medium. Both residues successfully generate cross-linked species in *E*. *coli* after UV-irradiation. These cross-links require the incorporation of the NNAA at specific positions; WT PRORP1 produces no cross-linked bands and T180 NNAA mutants lack the tRNA-sized shift in bands. Given the stability of PRORP1 in lysis buffer in the absence of reducing agents, this method should be successfully applied to pAzF-substituted PRORP under *in vitro* conditions, only requiring a slight modification to the purification and cross-linking buffers.

After proteinase K treatment, inhibition of RTase read-through by the cross-linked substrate is low, making definitive identification of the site of cross-linking difficult. Given that the cross-link should be irreversible (13, 15), high read-through is likely a result of the small PRORP1 peptide (SQ[pBF]HY) attached to pre-tRNA after proteolysis. Several potential solutions could decrease RTase read-through to determine a cross-linking site more definitively. Trypsin or another alternative protease could be used for the proteolysis step. For instance, trypsin would leave a much larger fragment (NGVQLSQ[pBF]HYNVLLYVCSLAEAATESSPNPGLSR) attached to the RNA. Likewise, partial digestion would leave larger fragments on average, which should decrease read-through. Excluding the proteolysis step entirely would likely lead to complete inhibition of read-through, but could also lead to poor resolution of the cross-linking site or no reaction with RTase. Furthermore, we re-purified the RNA by phenol-chloroform extraction and ethanol precipitation and un-proteolyzed protein-RNA would likely separate at the aqueous-organic interface instead of the aqueous phase. This change could be accommodated by precipitating/concentrating the sample without phenol-chloroform extraction. These changes will yield a cross-linking method that is more broadly-applicable.

In addition to the PRORP1 cross-linking we observe in the *E*. *coli* expression cells, the high read-through of the cross-linked RNA after complete proteinase K digestion suggests that this method might be adapted to identify the *in vivo* RNA binding partners and substrates of PRORPs. To achieve this, the RT step will be used to instead generate a cDNA library and the RNAs bound to the protein

identified using downstream RNA-seq methods. This could then be applied to human PRORP in tissue culture to identify its substrate pool. Data have been published that indicate PRORPs affect the processing of alternative precursor substrates (5, 6, 21), so an RNA-seq variation on our method would be useful for identifying the full population of RNA substrates with which the plant or human PRORPs interact. Protein-RNA cross-linking sites are frequently mutagenic in sequencing data (22), so the sequencing data could also be used to determine recognition patterns for the *in vivo* substrates.

The method we developed here will have broad use for determining the molecular interactions with RNA/substrates by protein-only RNase Ps and other RNA-binding proteins. For instance, the method could be used to test the predictive model we presented in Chapter 3. Furthermore, the cross-linking in *E. coli* cells indicates that we can pull out protein-RNA interactions using both cross-linkers tested. While we cannot yet definitively say whether the shift is a result of tRNA substrates, the molecular weight shift we observe is consistent with that being the case (20–25 kDa). This method could be further developed to include downstream RNA-seq methods to identify the RNA species bound, potentially expanding the known *in vivo* interaction partners for PRORPs or other RNA-binding proteins.

MATERIALS AND METHODS

Expression cell line preparation

We obtained the pEVOL plasmids via AddGene (Fig. C-5). These chloramphenicol resistant (CamR) plasmids encode one copy of the Amber suppressor tRNA ($tRNA^{SUP}$), as well as two copies each of the mutant *Methanocaldococcus jannaschii* aminoacyl-tRNA synthetases (*M.j*. muRS), which are responsible for charging $tRNA^{SUP}$ with either pAzF (14) or pBF (13). The genes are under the control of the arabinose operon promoter. In the presence of arabinose, araC no long binds to and inhibits the promoter, allowing transcription to initiate. We transformed the plasmids into BL21(DE3) *E*. *coli* and selected for transformants on LB/Cam agar plates.

BL21(DE3)/pEVOL cells were made Z-competent (Zymo Research, chemical competence kit T3002) following standard protocol. Briefly, colonies were selected from LB/Cam plates and grown overnight at 37°C in 5 mL LB/Cam. 0.5 mL of each overnight was added to 50 mL super optimal broth (SOB)/Cam and grown to OD_{600} ≈ 0.5, then transferred to ice. After ≥ 10 minutes, cells were pelleted at 4°C, then resuspended and washed once each with Wash and Competent buffers. Cells were aliquoted on ice, then flash frozen with liquid nitrogen and stored at -80°C until further transformation. Z-competent BL21(DE3)/pEVOL-pAzF or BL21(DE3)/ pEVOL-pBF cells were thawed on ice, then transformed with pET-M11/ Δ76PRORP1 containing an Amber mutation (Fig. C-6). Transformants were selected using the pEVOL CamR and the kanamycin resistance (KanR) marker on pET-M11 by growing on LB/Cam/Kan agar plates.

Colonies for each mutant were picked from LB/Cam/Kan plates and grown in LeMaster and Richards minimal medium (LRMM)/Kan/Cam (90% of the final expression volume) with pETM-11/*At*PRORP1 (KanR) and either pEVOL/pAzF or pEVOL/pBF (CamR). Induction media was prepared in LRMM with a 10X working concentration each of arabinose (final concentration = 0.2 mM), IPTG (F.C. = 0.7 mM), and either pAzF (Chem-Impex, F.C. = 1 mM) or pBF (Chem-Impex, F.C. =

0.2 mM). The pBF was used at a lower concentration than pAzF because it is more cytotoxic. Both pAzF and pBF were prepared by dissolving the solid to a stock concentration (500 mM pAzF or 200 mM pBF) in 1 M NaOH, because they are not soluble at the stock concentrations at neutral pH. Addition of each to the induction media was offset by adding an equal volume of 1 M HCl. Upon reaching $OD_{600} \approx 0.8$, temperature was reduced to 18°C and room temperature induction media was added (10% final volume). PRORP1 expression was induced overnight (16–18 hours). 50 mL test expressions included/excluded the IPTG, arabinose, and NNAA as needed. For expressions that excluded the NNAA, equal volumes of 1 M NaOH and 1 M HCl were added to the induction media to control for potential effects from higher ionic strength in downstream handling.

Enzyme preparation

pBF-substituted PRORP1 was purified as described previously for WT PRORP1 (7, 10, 11), but samples were kept shielded from light whenever possible. Briefly, cells were lysed using a microfluidizer, then pelleted at 4°C. Standard purification buffer conditions include 30 mM 3-morpholinopropane-1-sulfonic acid (MOPS) pH 7.8, 1 mM TCEP, 10% glycerol, and variable NaCl and imidazole concentrations depending on the column resin used. Supernatant was loaded onto Ni^{2+}-sepharose, washed, and then eluted with an imidazole gradient. Protein was dialyzed into low salt conditions overnight, then loaded onto SP-sepharose to remove RNA. The sepharose column was washed with low salt buffer until A_{260} dropped < 0.05, then protein was eluted using an NaCl gradient. The His_6-tag was removed by incubating with His_6-tagged TEV protease (> 1 mg per 40 mg PRORP1) overnight while dialyzing against buffer containing 30 mM imidazole at 4°C, which was then removed by re-loading onto the Ni^{2+}-sepharose column and washing with buffer containing 30 mM imidazole (PRORP1 collected in the flow-through). Protein was concentrated, then dialyzed into a size-exclusion buffer (30 mM MOPS pH 7.8, 300 mM NaCl, and 1 mM TCEP), and finally loaded onto a HiPrep 26/60 Sephacryl S-200 HR SEC column. Final eluted protein was concentrated, dialyzed against storage buffer (30 mM MOPS pH 7.8, 100 mM

NaCl, and 1 mM TCEP, prepared with CHELEX-treated Milli-Q (MQ) H_2O and metal-free components) containing 1 mM ethylenediaminetetraacetic acid (EDTA), then storage buffer without EDTA.

Purifying the pAzF-substituted PRORP1 under standard conditions resulted in a protein that did not have the expected mass-shift by Q-TOF MS (Fig. C-4). The normal buffer conditions contain TCEP, which can react with azides (Staudinger reaction, Scheme 4-1). DTT and βME can both reduce azides, although with different reaction kinetics. We incubated the azide-containing protein for seven days at 4°C in lysis buffers with or without reducing agent. The protein with no reducing agent resulted in small amounts of additional truncation products, but did not precipitate to a greater extent than the three reducing conditions (Fig. 4-8). The binding affinity and activity of these proteins were not measured.

Q-TOF Mass Spectrometry

Each injection on the Q-TOF used 8 µL of ≈ 10 µM PRORP1 to obtain a sufficient signal for deconvolution (data not shown). Samples were exchanged into MQ H_2O using 7 kDa MWCO Zeba™ Micro Spin Desalting Columns (Thermo Fisher; protocol as in instruction manual). Significant sample was lost during desalting (as determined by A_{280}), so a 20-25 µM sample was loaded onto the column.

Desalted proteins were loaded onto a Poroshell C8 column (Agilent) in 5% acetonitrile on a 1290 Infinity II liquid chromatography (LC) system coupled to a 6250 Accurate Mass Q-TOF LC/MS (Agilent). Proteins were washed for 5 minutes with the loading conditions, then eluted using a 15 minute linear gradient to 100% acetonitrile. PRORP1 variants typically elute 6-7 minutes into the gradient. Masses for individual peaks were determined by the intact-protein deconvolution program in the Agilent MassHunter Qualitative Analysis software.

RNA preparation

We prepared the *B. subtilis* pre-$tRNA^{Asp}$ substrate with a 5-nt 5’ leader and 3’ discriminator base by *in vitro* transcription as described previously (7). We used a

PCR-generated DNA template with a T7 RNA polymerase promoter and pre-tRNA with a 5-nt 5' leader and a GCCA 3' trailer. Fluorescein-labeled pre-tRNA was generated by transcribing with excess 5'-monophosphorothioate guanosine (GMPS) and reacting with 5-iodoacetamido fluorescein (5-IAF) at 37°C overnight. Substrates were gel-purified by 12% urea-PAGE, crushing the band into 45 mL TE-SDS and eluting overnight at 4°C.

Photo-activation for cross-linking

We cross-linked samples *in vitro* using the set-up described in Figure C-7. Samples were pre-incubated for ≥ 20 minutes in buffer containing Ca^{2+}, which does not activate PRORP1 catalysis. Samples were placed into the lids of Eppendorf tubes (tubes and lids were labeled with matching IDs), allowing us to re-collect the irradiated sample by centrifugation after replacing the tube. For the time-course experiment (Fig. 4-5A), 10 µL drops containing 400 nM substrate and 2 µM Y133pBF PRORP1 were irradiated at 365 nm for various times as listed. For the large-scale cross-linking experiment (Fig. 4-5B), 10 µL drops containing 10 µM substrate and 15 µM Y133pBF PRORP1 were irradiated at 365 nm for 120 minutes.

Samples were re-purified by denaturing urea-SDS-PAGE for use in primer extension assays. Irradiated substrate bands and cross-linked substrate bands were cut from the gel and crushed into 45 mL TE-SDS and eluted overnight at 4°C. Gel fragments were filtered and samples concentrated on 10 kDa MWCO Amicon® Ultra Centrifugal filters. Cross-linked substrate was treated with proteinase K (final concentration ≈ 50 µg/mL) for 1 hour at 37°C. The proteolyzed sample was then phenol-chloroform (isoamyl alcohol stabilized) extracted and ethanol precipitated. Irradiated substrate was also phenol-chloroform extracted and ethanol precipitated.

Primer extension

For our primer extension assays, we used the Omniscript RT kit (Qiagen). We used two fluorescein-labeled oligo-DNA primers (IDT) in the reverse transcription

reactions (Fig. 4-6). Each primer contained higher molecular weight bands, which ran at the same molecular weight as some of the primer extension bands, interfering with our ability to analyze the RT gels. Thus, primers were gel-purified using 12% urea-PAGE, overnight elution, Amicon concentration, and ethanol-precipitation. Primer 1 is optimal for observing cross-linking to the pre-tRNA between the TψC-loop and the anti-codon stem. Primer 2 is optimal for observing cross-linking to the regions between the D-arm and the 5' end of the pre-tRNA.

Templates were mixed with the oligo-DNA primers, then unfolded at 95°C for 3 minutes. The primed-RNA was cooled to 37°C for > 15 minutes, after which each RT reaction was initiated by the addition of a master mix. Master mix contained 2x each of Omniscript RT buffer, Omniscript RTase, dNTPs, and SUPERase•In (Thermo Fisher). Reactions were incubated at 37°C for 60 minutes, then loaded onto urea-PAGE (glycerol added for gel loading).

APPENDIX C

This appendix contains supporting figures for Chapter 4.

Figures

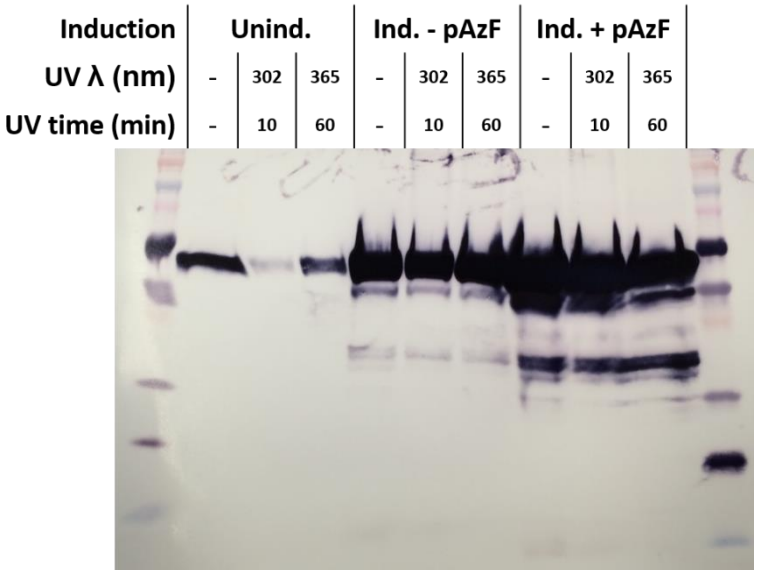


Figure C-1: Cross-linking with WT PRORP1 in *E. coli*. Protein was induced with or without pAzF to control for effects on cell viability. Regardless of irradiation conditions, no cross-linked band was observed as was the case for the NNAA-substituted proteins.

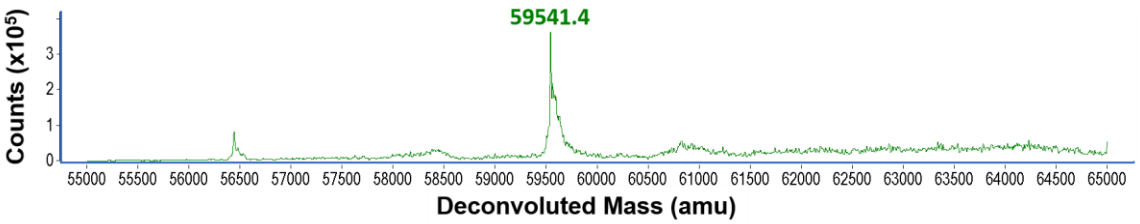


Figure C-2: Q-TOF MS deconvolution trace for Y133pBF His_6-TEV-Δ76PRORP1. Deconvolution range: 55000–65000 Da. Calculated WT mass (with His_6-tag) is 59450.8 Da. Calculated Y133pBF mass is 59538.9 (+88.1 Da). The observed mass (WT +90.6 Da) is consistent with successful pBF substitution.

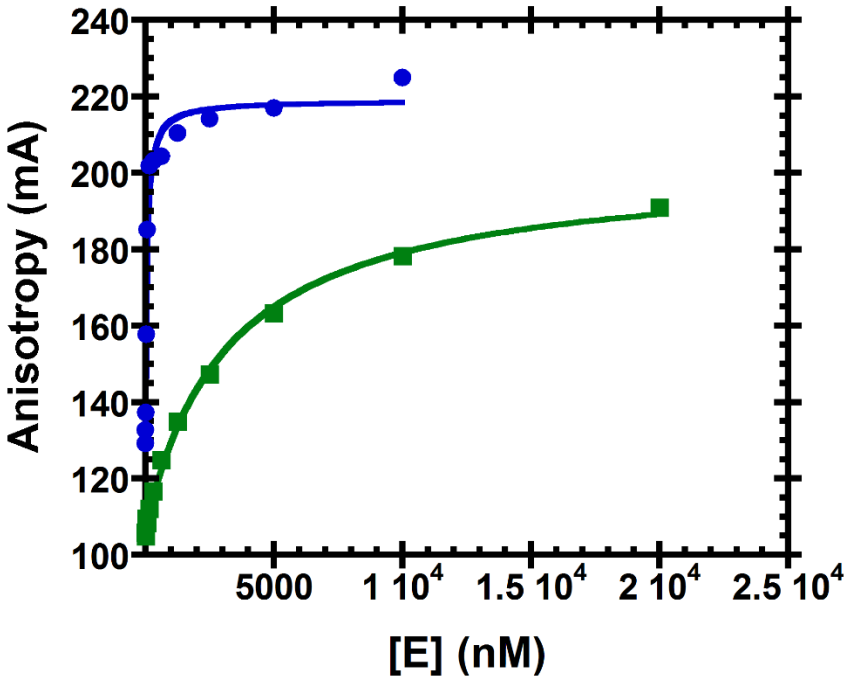


Figure C-3: Y133pBF-pre-tRNA binding curves. $K_{D,app}$ = 60 ± 10 nM (*B. subtilis* pre-tRNA^{Asp} (ptR5) in 100 mM NaCl), 3300 ± 300 nM (ptR5 in 330 mM NaCl). Y133pBF affinity for ptR5 is > 6-fold weaker than WT at 100 mM NaCl and ≈ 20-fold weaker than WT at 330 mM NaCl.

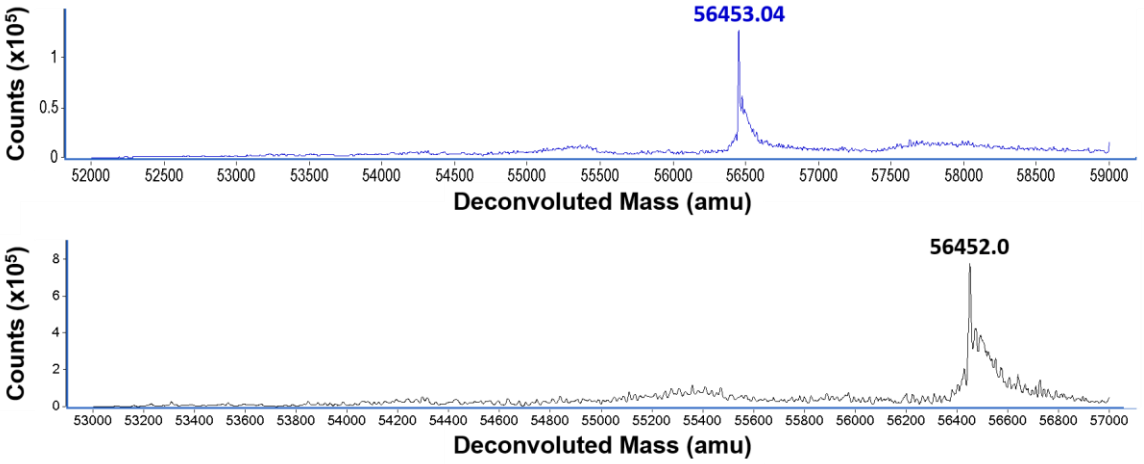


Figure C-4: Q-TOF MS deconvolution traces for WT (top) and Y133pAzF Δ76PRORP1 (bottom). (**Top**) Deconvolution range: 52000–59000 Da. Calculated WT mass after removal of the His_6-tag is 56452.5 Da. (**Bottom**) Deconvolution range: 53000–57000 Da (note the different scale). Calculated Y133pAzF mass is 56477.5 (+25 Da). The observed mass (WT -1 Da) is consistent with reduction of the azide to yield the amino-substituted protein.

Figure C-5: pEVOL plasmids encoding an Amber suppressor tRNA and *M. jannaschii* mutant tRNA synthetases (M.j. muRS) (13, 14). The plasmid also encodes a chloramphenicol resistance (CamR) marker and *araC*, the gene for the arabinose operon (araBAD) regulator protein. In the presence of arabinose, araC un-binds from the araBAD promoter, releasing it for transcription initiation.

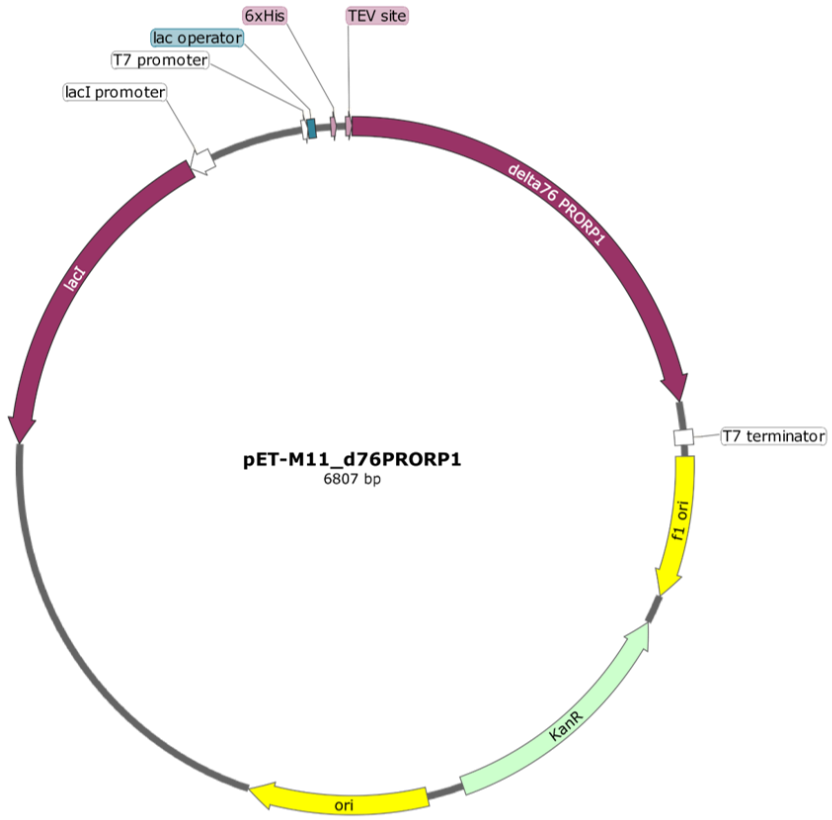


Figure C-6: pET-M11 plasmid encoding His_6-TEV-Δ76PRORP1 (11). The plasmid also encodes a kanamycin resistance (KanR) marker and *lacI*, the gene for the lactose operon repressor protein. In the presence of lactose (or an analog such as IPTG), lacI un-binds from the lac operator, releasing it for transcription initiation.

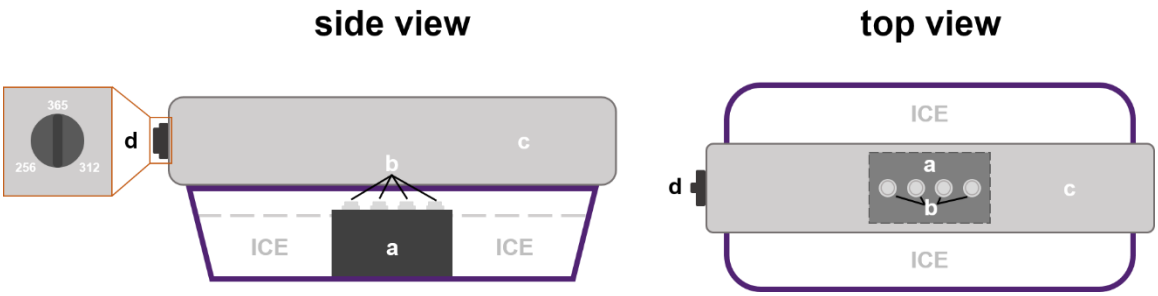


Figure C-7: UV cross-linking set-up. A heat-block (a) was placed upside down in an ice bucket (*purple*) and surrounded with ice. Samples were placed in Eppendorf tube lids (b) upside down on the heat block. The UV lamp (c) was rested ≈ 0.5" above the samples and the dial (d) was set to the λ required for a given protocol.

REFERENCES

1. Klemm BP, Wu N, Chen Y, Liu X, Kaitany KJ, Howard MJ, Fierke CA. The Diversity of Ribonuclease P: Protein and RNA Catalysts with Analogous Biological Functions. *Biomolecules*. 2016;6(2):E27.
2. Walker SC, Engelke DR. Ribonuclease P: the evolution of an ancient RNA enzyme. *Crit Rev Biochem Mol Biol*. 2006;41(2):77-102.
3. Lechner M, Rossmanith W, Hartmann RK, Thölken C, Gutmann B, Giegé P, Gobert A. Distribution of Ribonucleoprotein and Protein-Only RNase P in Eukarya. *Mol Biol Evol*. 2015;32(12):3186-93.
4. Holzmann J, Frank P, Löffler E, Bennett KL, Gerner C, Rossmanith W. RNase P without RNA: identification and functional reconstitution of the human mitochondrial tRNA processing enzyme. *Cell*. 2008;135(3):462-74.
5. Gobert A, Gutmann B, Taschner A, Gößringer M, Holzmann J, Hartmann RK, Rossmanith W, Giegé P. A single *Arabidopsis* organellar protein has RNase P activity. *Nat Struct Mol Biol*. 2010;17(6):740-4.
6. Gutmann B, Gobert A, Giegé P. PRORP proteins support RNase P activity in both organelles and the nucleus in *Arabidopsis*. *Genes Dev*. 2012;26(10):1022-7.
7. Howard MJ, Klemm BP, Fierke CA. Mechanistic Studies Reveal Similar Catalytic Strategies for Phosphodiester Bond Hydrolysis by Protein-only and RNA-dependent Ribonuclease P. *J Biol Chem*. 2015;290(21):13454-64.
8. Gobert A, Pinker F, Fuchsbauer O, Gutmann B, Boutin R, Roblin P, Sauter C, Giegé P. Structural insights into protein-only RNase P complexed with tRNA. *Nat Commun*. 2013;4(1353):1-8.
9. Brillante N, Gößringer M, Lindenhofer D, Toth U, Rossmanith W, Hartmann RK. Substrate recognition and cleavage-site selection by a single-subunit protein-only RNase P. *Nucleic Acids Res*. 2016;44(5):2323-36.
10. Howard MJ, Karasik A, Klemm BP, Mei C, Shanmuganathan A, Fierke CA, Koutmos M. Differential substrate recognition by isozymes of plant protein-only Ribonuclease P. *RNA*. 2016;22(5):782-92.
11. Howard MJ, Lim WH, Fierke CA, Koutmos M. Mitochondrial ribonuclease P structure provides insight into the evolution of catalytic strategies for precursor-tRNA 5 ' processing. *Proc Nat Acad Sci U S A*. 2012;109(40):16149-54.
12. Imai T, Nakamura T, Maeda T, Nakayama K, Gao X, Nakashima T, Kakuta Y, Kimura M. Pentatricopeptide repeat motifs in the processing enzyme PRORP1 in *Arabidopsis thaliana* play a crucial role in recognition of nucleotide bases at TψC loop in precursor tRNAs. *Biochem Biophys Res Commun*. 2014;450(4):1541-6.
13. Chin JW, Martin AB, King DS, Wang L, Schultz PG. Addition of a photocrosslinking amino acid to the genetic code of *Escherichia coli*. *Proc Nat Acad Sci U S A*. 2002;99(17):11020-4.
14. Chin JW, Santoro SW, Martin AB, King DS, Wang L, Schultz PG. Addition of *p*-Azido-L-phenylalanine to the Genetic Code of *Escherichia coli*. *J Am Chem Soc*. 2002;124(31):9026-7.

15. Lee HS, Dimla RD, Schultz PG. Protein-DNA photo-crosslinking with a genetically encoded benzophenone-containing amino acid. *Bioorg Med Chem Lett*. 2009;19(17):5222-4.
16. LeMaster DM, Richards FM. Preparative-scale isolation of isotopically labeled amino acids. *Anal Biochem*. 1982;122:238-47.
17. Jackman JE, Montange RK, Malik HS, Phizicky EM. Identification of the yeast gene encoding the tRNA m^1G methyltransferase responsible for modification at position 9. *RNA*. 2003;9(5):574-85.
18. Staros JV, Bayley H, Standring DN, Knowles JR. Reduction of aryl azides by thiols: implications for the use of photoaffinity reagents. *Biochem Biophys Res Commun*. 1978;80(3):568-72.
19. Loria A, Niranjanakumari S, Fierke CA, Pan T. Recognition of a pre-tRNA substrate by the *Bacillus subtilis* RNase P holoenzyme. *Biochemistry*. 1998;37(44):15466-73.
20. Niranjanakumari S, Day-Storms JJ, Ahmed M, Hsieh J, Zahler NH, Venters RA, Fierke CA. Probing the architecture of the *B*. *subtilis* RNase P Holoenzyme active site by crosslinking and affinity cleavage. *RNA*. 2007;13:512-35.
21. Rackham O, Shearwood AM, Mercer TR, Davies SM, Mattick JS, Filipovska A. Long noncoding RNAs are generated from the mitochondrial genome and regulated by nuclear-encoded proteins. *RNA*. 2011;17(12):2085-93.
22. Zhang C, Darnell RB. Mapping *in vivo* protein-RNA interactions at single-nucleotide resolution from HITS-CLIP data. *Nature Biotech*. 2011;29(7):607-14.

CHAPTER 5

Conclusions and Future Directions

In this thesis, we have determined several key features of substrate recognition by plant PRORPs. We identified both the minimal pre-tRNA structural requirements for recognition by PRORP and a novel RNA recognition surface on PRORPs that is important for substrate binding. We identified an alternative 5' end selection for PRORPs that results in miscleavage products *in vitro*. We also characterized the mode of substrate recognition by varying ionic strength to parse the ionic and non-ionic components of the binding affinity. Finally, we made progress toward developing a method for mapping the PRORP-substrate complex.

CONCLUSIONS

Structural features of pre-tRNA recognized by *At*PRORPs

We characterized several features of pre-tRNAs that are recognized by PRORPs to compare them with the well-studied RNA-dependent RNase Ps. The RNA-dependent RNase Ps utilize a variety of different pre-tRNA structural features for recognition. Conserved components of the bacterial RNase P RNA specificity domain recognize the tRNA D- and TψC-loop structures (1), including stacking interactions with a conserved long range G:C base pair and sugar face interactions with the conserved terminal G:C base pair of the TψC-stem. Like the RNA-dependent RNase P, PRORPs require a D-loop for tight binding and efficient catalysis (Fig. 2-3, Table A-2) (2, 3).

The bacterial RNase P RNA makes specific interactions with the 3' sequence and the RNase P protein interacts with the 5' leader beyond 2-nt (4, 5). In contrast,

PRORPs do not require a 3’ trailer sequence and can efficiently recognize pre-tRNA substrates with 5’ leaders containing only 1- or 2-nt (Table 2-1) (2, 3). These data are consistent with the alterations of slope and intercept for the NaCl-dependence of affinities for pre-tRNA substrates containing various leader lengths (Fig. 3-4).

*At*PRORP substrate specificity

The substrate recognition strategies for each of the three *At*PRORP isozymes are similar. The catalytic efficiencies for a given *At*PRORP with the four pre-tRNA substrates varied < 6-fold (Fig. 2-5) (3). Although the binding affinity for various pre-tRNA substrates differs by up to 100-fold, the three *At*PRORPs displayed affinities for a given substrate that varied < 4-fold (Fig. 2-4) (3). Additional substrates should be screened, for instance with a high-throughput method such as HITS–KIN before the lack of substrate selectivity we observe for *At*PRORPs can be extended (6), because the uniform catalytic efficiencies we observe may not hold for all substrates in the much larger population of plant pre-tRNAs.

Alternative 5’ end selection by *At*PRORPs

PRORPs display a 5’ end selection that is alternative to the RNA-based RNase P (Fig. 2-6) (3). *At*PRORPs produce an N_{-2}/N_{-1} miscleavage product with various pre-tRNAs *in vitro*, yet *B*. *subtilis* RNase P produces only the correct product with the substrates we tested (3). All substrates that display a miscleavage pattern have the potential for N_{-1}/N_{73} base pairing (U:A, A:U, or C:G; Fig. 2-3, Table 5-1) (3, 7). The bacterial RNase P makes several interactions that could improve fidelity for these substrates. For instance, the RNA component base-pairs with uridines at the N_{-1} position and with the 3′ discriminator base of pre-tRNA, which could increase selection for the correct 5′ product (1, 8). This is consistent with the minor miscleavage we observed for the *B. subtilis* RNase P-catalyzed cleavage of a chloroplast pre-$tRNA^{Phe}$, which instead has a cytidine at N_{-1} (Fig. 2-6). Thus, PRORPs use different 5’ end recognition strategies than bacterial RNase P.

Table 5-1: Miscleavage observed for *At*PRORPs or *B*. *subtilis* RNase P.[a]

Substrate	N_{-1}/N_{73}	N_{-1} Micleavage observed?	
		*At*PRORP	*B*. *subtilis* RNase P
B. *subtilis* Asp	U:G	no	no
A. thaliana (chlor) Phe	C:A	yes	yes
A. thaliana (nuc) Cys	A:U	yes	no
A. thaliana (mito) Cys	U:A	yes	no
A. thaliana (nuc) Phe	U:A	yes	no
Potato (mito) His[b]	G:C	yes	ND
E. *coli* His[c]	G:C	yes	yes

a: Unless otherwise denoted, data are from Chapter 2 (3).
b: From Placido, *et al*. 2010 (7).
c: From Brillante, *et al*. 2016 (2).

We observe increased 5’ end selection fidelity for *At*PRORP1 when compared with either *At*PRORP2 or *At*PRORP3. This suggests that *At*PRORP1 makes different interactions at the 5’ end of pre-tRNA. In fact, our PRORP-substrate model indicates that a loop in the NYN domain – for which *At*PRORP2/3 has a four-residue insertion – is positioned to recognize the pre-tRNA 5’/3’ ends (Fig. 3-7). *At*PRORPs can reprocess the miscleaved 1-nt pre-tRNA product (Fig. A-3), suggesting that the alternative cleavage site may not lead to an accumulation of mis-processed tRNAs, which would be detrimental to tRNA function *in vivo*.

*At*PRORP-substrate molecular recognition

We characterized the PRORP-substrate complex with respect to the salt dependence to parse the contributions of ionic interactions to the complex affinity. These data indicate that *At*PRORPs utilize 4–5 direct contacts to backbone phosphate for substrate recognition (Fig. 3-2C, Table 3-1). At least one of these contacts is located within the 5’ leader. The steric restrictions on the complex that are imposed by the required placement of the 5’ end at the active site and the requirement for contacts to the D- and TψC-loops allow us to make further predictions about the recognition complex. For instance, the backbone contacts within the tRNA body are likely located within the acceptor- or TψC-stems, or the D- or TψC-loops.

Our data for NaCl-inhibition of the STO k_{obs} are consistent with an *At*PRORP kinetic recognition mechanism that includes steps in addition to the simple

mechanism previously proposed (Fig. 3-6, Scheme 3-1). The data impose one of two restrictions on the recognition mechanism: 1) the rate constant for unbinding (k_{off}) must be equal to or less than the rate constant for chemistry (k_{chem}), such that the binding is not in rapid-equilibrium, or 2) there are additional kinetic steps beyond substrate binding such as chemistry that are affected by NaCl. The latter case would be consistent with *At*PRORP kinetic pH-dependence data. The pH dependence of the single-turnover k_{obs} for cleavage of two pre-tRNA substrates catalyzed by *At*PRORP is consistent with a single ionization with maximal activity at high-pH (9), however, the apparent pK_a values are too low to be pK_a for the Mg•H_2O ionization. Thus, the pH-dependence plateau, which yields different maximal k_{obs} values for the two substrates, likely represents a pH-independent step such as the binding event or a subsequent non-hydrolytic kinetic step. Further, we observe biphasic single-turnover kinetics with the *Arabidopsis* mitochondrial pre-$tRNA^{Cys}$ substrate at high pH, with the slower phase being pH-independent.

These data are consistent with a PRORP kinetic mechanism that includes a step after binding whose rate is dependent on the concentration of NaCl but not on the pH (Scheme 3-2). This step could involve any number of kinetic recognition methods, such as repositioning the substrate within the active site, separating the 5' and 3' ends, conformational changes within the protein, or a combination of these. My working hypothesis for these observations is that the additional kinetic recognition step involves repositioning the substrate on the NYN domain, including repositioning helix α21 and the repositioning the 5' leader near the α16–α17 loop. This model could explain the kinetic defects we observe in His 498 variants (9), as well as the differences in miscleavage we observe for PRORP1 and PRORP2/3 (Fig. 2-6).

*At*PRORP PPR domain

We identified several residues in the *At*PRORP1 and *At*PRORP2 PPR domains that are important for substrate binding (Chapter 3). A broad screen of alanine mutants on the putative substrate binding surface of *At*PRORP2 identified several residues that are important for binding and/or protein stability (Fig. 3-5A, Table B-

2). We characterized this surface further with *At*PRORP1 and determined the types of interactions each residue makes with substrates (Fig. 3-5B, Table 3-2). Using this analysis, we determined that conserved tyrosine (*At*PRORP1 Tyr 140) and arginine (*At*PRORP1 Arg 184) residues that are critical for substrate recognition. The data for Y140F and Y140A indicate that the tyrosine side chain makes both stacking interactions and a hydrogen bond with substrate. Likewise, binding data with R184A and R184K suggests that the guanidinium group makes one or more interactions with substrate. It remains possible that the binding defect in the R184K variant is due to the side chain being incorrectly positioned for substrate interaction. These residues are completely conserved in PRORPs from plants to humans and are likely a common feature that is critical for substrate recognition. Furthermore, these residues are at positions that are not regularly conserved in other PPR-containing proteins. Importantly, the data are consistent with substrate recognition by *At*PRORPs that does not rely on the same nucleobase-specifying interaction modes as PPR domains from single-stranded RNA-binding proteins.

Methods to map PRORP-substrate complex

We developed a PRORP-substrate cross-linking method which we can use to map the complex (Chapter 4). The method includes a work-flow which utilizes non-natural amino acid substitutions in PRORPs (Fig. 4-1, Fig. 4-2), each of which is photo-activatable and can cross-link to pre-tRNA substrates. Using a reverse transcription primer extension assay, we observe decreased read-through in the D-loop and acceptor stem when NNAAs are incorporated in *At*PRORP1 at Tyr 133 (Fig. 4-6). However, the Y133pBpF variant had significantly reduced substrate binding affinity (Fig. C-3), indicating that the mutated side chain affected the PRORP-substrate recognition complex.

FUTURE DIRECTIONS

Substrate recognition

There are still several outstanding questions about PRORP-substrate recognition. For instance, our single-turnover NaCl inhibition data indicate that the

kinetic mechanism for recognition likely involves steps in addition to the biomolecular binding event (Chapter 3). These steps could be interrogated using stopped-flow to measure the effects of NaCl on various kinetic steps. We might be able to make use of the *Arabidopsis* pre-tRNA^{Asp} substrate, with which we observe biphasic kinetics under high pH single-turnover conditions (9). We proposed that this may be the result of a second population of substrate conformer, but an alternative hypothesis could involve another kinetic recognition step that becomes rate-limiting under high-pH conditions. We could test this hypothesis by varying the NaCl concentrations and quantifying the effects on the second phase, both by stopped-flow and the gel-based assay.

Another outstanding question regards the 5' end selection (Chapter 2) (3). We observe partial miscleavage in substrates that contain a U:A base pair at N_{-1}/N_{73} and a reduction in the fraction miscleaved for a substrate containing A:U at N_{-1}/N_{73} when compared to substrates containing U:A (Fig. 2-6) (3). However, we also observe differences in the fraction miscleaved for two substrates containing U:A at N_{-1}/N_{73}. Thus, there are likely to be additional pre-tRNA structural determinants for the miscleavage pattern.

Given that we observe different miscleavage patterns for different pre-tRNAs with the same N_{-1}/N_{73} pair, work needs to be done to vary the N_{-1}/N_{73} identities within the same substrate. In work published by Brillante, *et al*., the variation of the N_{-1}/N_{73} pair to G–C yielded nearly 100% miscleavage, while U–A yielded both the correct and miscleavage products (2). A more complete set of sequence variations should include identities (i.e. A:A, U:U, C:C, and G:G), purine-purine pairs (A:G and G:A), and pyrimidine pairs (U:C and C:U) in addition to various pairs with base-pairing potential (U:A, A:U, C:G, G:C, U:G, G:U, etc.). If we determine whether the miscleavage correlates primarily with the strength of the potential base pair, or with the positions of the purine and pyrimidine for a given base pair, the data will yield valuable information about the 5' end selection requirements. This will also allow us to identify substrates that are miscleaved *in vivo*, as is the case with the potato mitochondrial pre-tRNA^{His} (7).

Furthermore, the molecular features of the D- and TψC-loops that are required for PRORP recognition remain to be determined (Chapter 2) (3). For instance, Brillante, *et al*. attempted to vary the sequence at several positions in the D- and TψC-loops (2). Alterations to the D-loop yielded minimal changes, while variation within the TψC-loop only affected binding and catalysis in a substrate that lacked the D-loop (2). Likewise, we mutated four uridines in the D-loop that are not making structurally important contacts to adenosine in *B*. *subtilis* pre-tRNA^{Asp}, each of which affected the *At*PRORP1 substrate affinity < 2-fold. It might be the case that many or all contacts that PRORPs make to the D- and TψC-loops are stacking/van der Waals interactions or hydrogen bonding to the backbone, which are not expected to be sequence-specific. We could test various types of non-specific interactions using substrates with single-atom substitutions, such as phosphorothioate substitutions at non-bridging oxygens (R_P- and S_P-phosphorothioate) and 2' modifications in the sugar (2'-NH_2, 2'-F, 2'-H etc.).

Finally, we proposed a binding model that leads to several predictions for the recognition complex (Chapter 3). The model suggests that the NYN helix α21 is not positioned for substrate binding in the substrate-free crystal structures. This is consistent with single-turnover data for *At*PRORP1 mutant His 498, which displays reduced catalysis (9). Further, the model suggests that the NYN domain α16–α17 loop is involved in 5' end recognition. This is consistent with the observation that *At*PRORP2–3 have both a different loop sequence than *At*PRORP1 and a higher fraction of alternative 5' end selection. We can test both hypotheses using a variety of experiments. The cross-linking method developed in Chapter 4 could be used to identify the parts of substrate that contact the NYN domain. The *At*PRORP1 and *At*PRORP3 α17–α18 loop sequences could be swapped. The loop-swap variants should display the 5' end selection of the other *At*PRORP for a given substrate.

PRORP-substrate complex mapping

Going forward, we should take care to screen any NNAA-substituted variants for substrate binding affinity prior to cross-linking. We might be able to position the NNAA substitutions near, but not at, the binding surface to both cross-link the

substrate and reduce the impact on substrate binding affinity. The cross-linking method we developed will require modifications and further optimization to yield a more high-throughput approach (Chapter 4). For instance, digesting the protein with proteases that leave larger peptide fragments on the pre-tRNA substrates than does proteinase K could yield more efficient reductions to the primer extension at sites of cross-linking, which will yield more definitive information about the sites of contact. Furthermore, due to high read-through when proteinase K is used, we will be able append an RNA-seq experiment to the method (10), thus allowing us to pull RNA substrates and binding partners for various PRORP out of cells and identify them. The cross-linking is frequently mutagenic, so that cross-linking sites for any novel RNAs can be identified (11). These variations on the method we developed will yield large data sets from which we can gain valuable information about 1) the structure and dynamics of the PRORP-substrate complex and 2) the RNAs that PRORPs interact with *in vivo*, including non-pre-tRNAs that have not yet been identified as PRORP substrates.

Human PRORP-substrate complex

The metazoan PRORPs, such as those in humans, have additional complexity in the PRORP-substrate recognition complex. Rather than one protein and one RNA, as is the case in plant PRORPs, the human enzyme contains three distinct proteins, with a minimum six subunits, and at least one RNA substrate in the catalytically-competent complex (12). A sub-complex of the two subunits not utilized by the plant PRORPs acts a tRNA/pre-tRNA methyltransferase (13). This three-protein complex is responsible for 5’ end maturation for pre-tRNAs encoded by the human mitochondrial genome, yet little is known about the functional reason for the additional subunits. It has been hypothesized that the requirement stems from the non-canonical secondary and tertiary structures formed by the metazoan mitochondrial tRNAs (12, 13). It might be the case that these structures cannot be recognized by human PRORP using the same strategies defined in this thesis, thus requiring the additional subunits to form the correct substrate. This could be by altering the structure through methylation, as is achieved for mitochondrial

$tRNA^{Lys}$ (14), or by the substrate-methyltransferase sub-complex acting as a co-substrate for human PRORP. In either case, it is unlikely that the precise mode of PRORP-substrate recognition defined for plant PRORPs will be employed by human PRORP. Thus, the human PRORP-substrate complex will need to be defined. The cross-linking method in Chapter 4 could be useful to map the protein-RNA interactions for each of the three protein components, as well as to map the protein-protein interactions by combining the cross-linking with mass-spec.

OUTLOOK

On publishing this thesis, it has been 46 years since RNase P activity was detected in *E*. *coli* extracts (15), 34 years since a bacterial RNase P was shown to be an RNA catalyst (16), and nine years since a protein-only RNase P was purified from human mitochondria (12). Through the duration of this field, there have been many advancements to our understanding of catalytic strategies employed by nucleases. RNase Ps present us with the opportunity to study two independently evolved sets of enzymes that have distinct macromolecular composition, yet perform an identical biological function. This extreme case of convergent evolution continues to yield exciting new directions for biological, chemical, and biochemical research.

REFERENCES

1. Reiter NJ, Osterman A, Torres-Larios A, Swinger KK, Pan T, Mondragón A. Structure of a bacterial ribonuclease P holoenzyme in complex with tRNA. *Nature*. 2010;468(7325):784-9.
2. Brillante N, Gößringer M, Lindenhofer D, Toth U, Rossmanith W, Hartmann RK. Substrate recognition and cleavage-site selection by a single-subunit protein-only RNase P. *Nucleic Acids Res*. 2016;44(5):2323-36.
3. Howard MJ, Karasik A, Klemm BP, Mei C, Shanmuganathan A, Fierke CA, Koutmos M. Differential substrate recognition by isozymes of plant protein-only Ribonuclease P. *RNA*. 2016;22(5):782-92.
4. Crary SM, Niranjanakumari S, Fierke CA. The protein component of *Bacillus subtilis* ribonuclease P increases catalytic efficiency by enhancing interactions with the 5' leader sequence of pre-tRNA^{Asp}. *Biochemistry*. 1998;37(26):9409-16.
5. Rueda D, Hsieh J, Day-Storms JJ, Fierke CA, Walter NG. The 5' leader of precursor tRNA^{Asp} bound to the *Bacillus subtilis* RNase P holoenzyme has an extended conformation. *Biochemistry*. 2005;44(49):16130-9.
6. Guenther U-P, Yandek LE, Niland CN, Campbell FE, Anderson D, Anderson VE, Harris ME, Jankowsky E. Hidden specificity in an apparently nonspecific RNA-binding protein. *Nature*. 2013;502(7471):385-8.
7. Placido A, Sieber F, Gobert A, Gallerani R, Giegé P, Maréchal-Drouard L. Plant mitochondria use two pathways for the biogenesis of tRNA^{His}. *Nucleic Acids Res*. 2010;38(21):7711-7.
8. Zahler NH, Sun L, Christian EL, Harris ME. The pre-tRNA nucleotide base and 2'-hydroxyl at N(-1) contribute to fidelity in tRNA processing by RNase P. *J Mol Biol*. 2005;345(5):969-85.
9. Howard MJ, Klemm BP, Fierke CA. Mechanistic Studies Reveal Similar Catalytic Strategies for Phosphodiester Bond Hydrolysis by Protein-only and RNA-dependent Ribonuclease P. *J Biol Chem*. 2015;290(21):13454-64.
10. Ule J, Jensen KB, Ruggiu M, Mele A, Ule A, Darnell RB. CLIP Identifies Nova-Regulated RNA Networks in the Brain. *Science*. 2003;302(5648):1212-15.
11. Zhang C, Darnell RB. Mapping *in vivo* protein-RNA interactions at single-nucleotide resolution from HITS-CLIP data. *Nature Biotech*. 2011;29(7):607-14.
12. Holzmann J, Frank P, Löffler E, Bennett KL, Gerner C, Rossmanith W. RNase P without RNA: identification and functional reconstitution of the human mitochondrial tRNA processing enzyme. *Cell*. 2008;135(3):462-74.
13. Vilardo E, Nachbagauer C, Buzet A, Taschner A, Holzmann J, Rossmanith W. A subcomplex of human mitochondrial RNase P is a bifunctional methlytransferase–extensive moonlighting in mitochondrial tRNA biogenesis. *Nucleic Acids Res*. 2012;40(22):11583-93.
14. Helm M, Giegé R, Florentz C. A Watson-Crick Base-Pair-Disrupting Methyl Group (m^1A9) Is Sufficient for Cloverleaf Folding of Human Mitochondrial tRNA^{Lys}. *Biochemistry*. 1999;38(40):13338-46.
15. Altman S, Smith JD. Tyrosine tRNA Precursor Molecule Polynucleotide Sequence. *Nature*. 1971;233(36):35-9.

16. Guerrier-Takada C, Gardiner K, Marsh T, Pace NR, Altman S. The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell*. 1983;35(3 Pt 2):849-57.

APPENDIX D

Contributions to studies into the catalytic mechanism of *At*PRORP1[†]

This work contributed to a publication containing mechanistic data that revealed similarities in the catalytic strategies of protein-only and RNA-dependent RNase Ps (1). To study the pH dependence of *Arabidopsis thaliana* PRORP1 and variants, we devised a buffering system that allowed us to maintain constant buffer concentrations and easily control changes in the total ionic strength. We applied this system, as well as our anisotropy binding assays, to study the effects of mutation to two histidine residues near the active site. We used our gel-based assay to establish and quantify the importance of the four fully-conserved aspartate residues for catalysis in *At*PRORP1.

Buffering system for PRORP pH-dependence experiments

The buffers that we selected to study the pH-dependence of *At*PRORP1 were 2-(*N*-morpholino)ethanesulfonic acid (MES, $pK_{a,25°C}$ = 6.1), tris(hydroymethyl) aminomethane (Tris, $pK_{a,25°C}$ = 8.07), and 2-amino-2-methyl-propanol (AMP, $pK_{a,25°C}$ = 9.7) (Fig. D-1). Each buffer was selected from Good's buffers (2, 3). These three buffers maintain buffering capacity with overlapping ranges from pH 5.5–10.5, which fully encompasses the range of pH that we tested (Fig. D-2).

[†]This work is adapted and expanded from reference: Howard, MH; Klemm, BP; Fierke, CA. Mechanistic studies reveal similar catalytic strategies for phosphodiester bond hydrolysis by protein-only and RNA-dependent Ribonuclease P. *J Biol Chem* 2015, 290(21):13454-64.

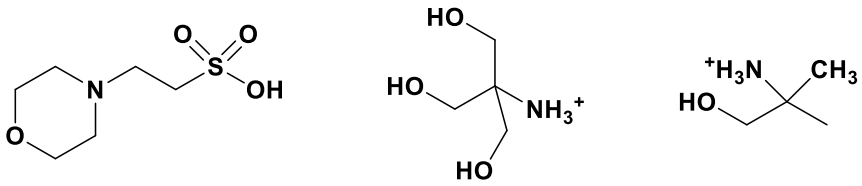


Figure D-1: The conjugate acids for each buffer used in this work. **Left**: 2-(N-morpholino)ethanesulfonic acid (MES). **Middle**: tris(hydroymethyl)aminomethane (Tris). **Right**: 2-amino-2-methyl-propanol (AMP).

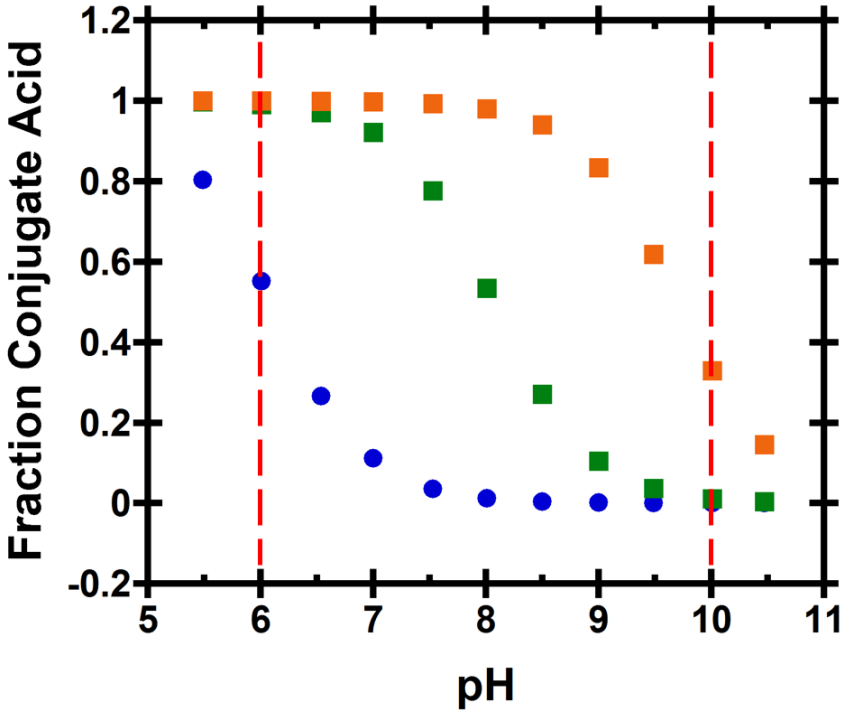


Figure D-2: The fraction of each buffer as the conjugate acid for every 0.5 pH step from 5.5 – 10.5, as calculated using the Henderson-Hasselbalch equation. The buffering range used in work presented is indicated by the dashed red lines. Buffers used are MES (blue), Tris (green), and AMP (orange). The conjugate acids may be neutral (circles) or positively charged (squares).

The ionic strength (*I*) of each buffer was calculated using equation 1, in which *c* is the concentration of the ion and *z* is the ion's charge. We adjusted to a constant total ionic strength using excess NaCl (Fig. D-3). At the time, we did not consider the NaCl concentration inhibition of catalysis (Chapter 3). However, the total NaCl concentration for these assays at the highest-NaCl condition (pH 8.5, < 200 mM) remains below the concentrations that significantly inhibited activity (> 400 mM).

$$I = \frac{1}{2}\sum_{i=1}^{n} c_i z_i^2 \quad (1)$$

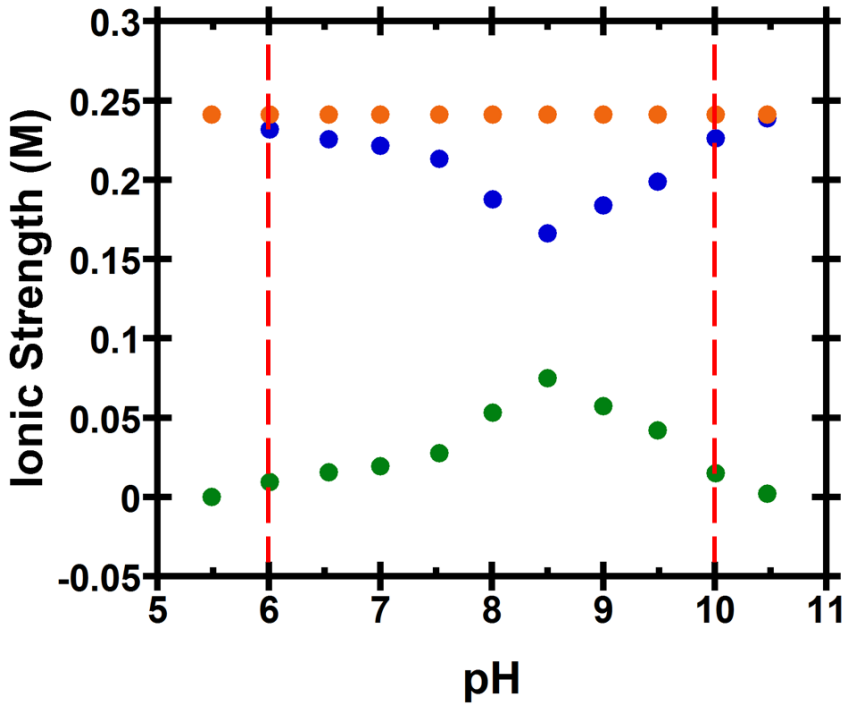


Figure D-3: Total ionic strength for our buffering system at 100 mM NaCl and 1 mM $MgCl_2$ (blue). The ionic strength was normalized to the maximum ionic strength (orange) by adding various amounts of excess NaCl (green). The buffering range used in work presented is indicated by the dashed red lines.

Histidine mutants' defects in catalysis and alterations in binding

Unlike the RNA-dependent RNase Ps, PRORPs have side chains that are capable of acid-base catalysis under physiological conditions. We identified two highly conserved histidine residues that are near the active site. His 438 is conserved in plants, while metazoans maintain a histidine one turn up on the same helix (His 447 in humans). His 498 is fully conserved in PRORPs. The H498A mutations reduced the observed single-turnover rate constant (k_{obs}) by > 80-fold compared to WT *At*PRORP1. The H498Q mutant had activity that was only ~ 4-fold reduced. These data are inconsistent with the residue participating in acid-base catalysis. The pH-dependencies of H438A and H498Q mutants revealed slight defects in catalysis, but only small changes to the observed pK_a (Fig. D-4). Equation 2 was fit to the data.

$$k_{obs} = \frac{k_{max}^{pH}}{1+10^{(pK_a - pH)}} \quad (2)$$

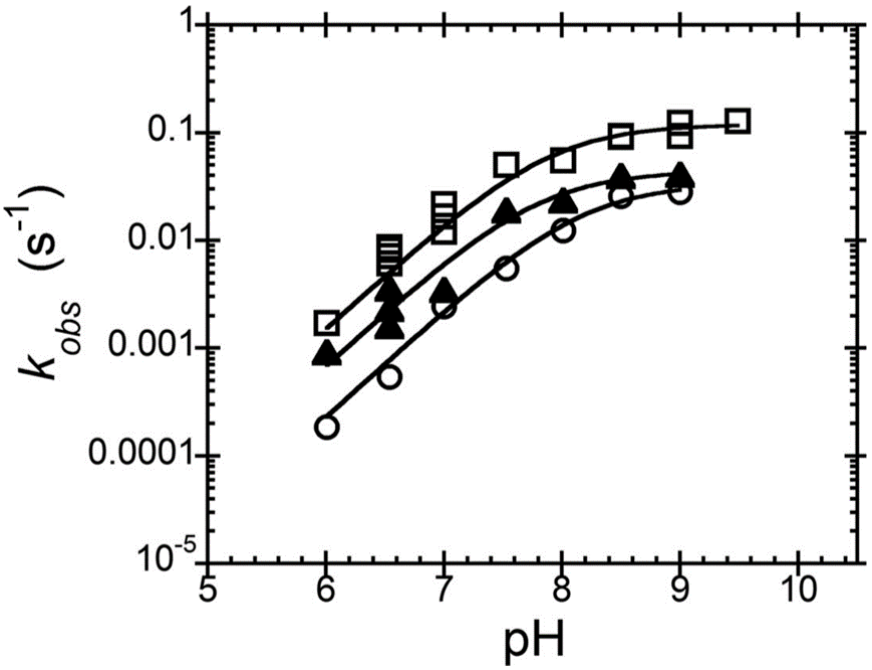


Figure D-4: pH-dependence of histidine mutants for cleavage of *A. thaliana* mitochondrial pre-tRNA^{Cys}. WT (open squares), H438A (triangles), and H498Q (open circles) are shown. The reactions are carried out using single turnover conditions, 5 μM enzyme and 20 nM substrate with the three component buffer. The k_{max} values are 0.12 ± 0.01, 0.03 ± 0.002, and 0.04 ± 0.003 s^{-1} for WT, H498Q, and H438A, respectively. The p$K_{a,app}$ values are 7.9 ± 0.1, 8.2 ± 0.1, and 7.8 ± 0.1 for WT, H498Q, and H438A, respectively.

The mutations did not affect the observed binding affinity (Fig. D-5). Equation 3 was fit to the data. The observed dissociation constants (K_D) are comparable to WT at ~ 700 nM. Interestingly, the endpoint anisotropy for both His 498 mutants was different, indicating potential differences in the rotational freedom of the fluorescein in the ES complex. We proposed that this is because His 498 is involved with positioning the tRNA and 5' leader on the complex.

$$A = A_0 + \frac{\Delta A[P]}{[P]+K_D} \quad (3)$$

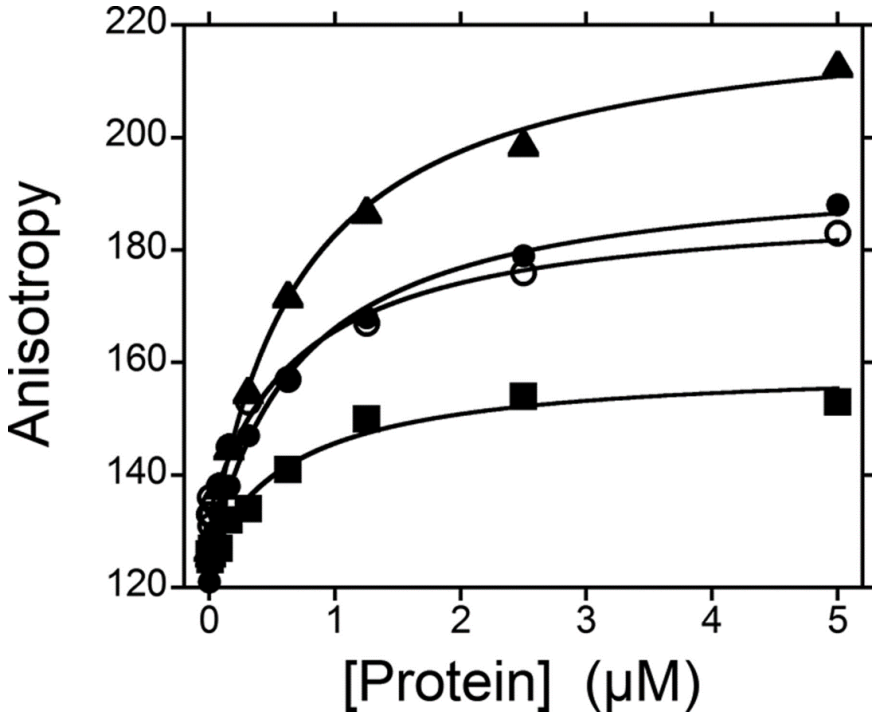


Figure D-5: Effects of H438A and H498A mutants on binding. Anisotropy binding curves for WT (circles), H498A (squares), H498Q (triangles), and H438A (open circles). Dissociation constants (K_D) are 700 ± 80, 680 ± 150, 710 ± 60, and 700 ± 120 nM for WT, H498A, H498Q, and H438A, respectively. Assay buffers included 20 mM MOPS pH 7.8, 300 mM NaCl, 1 mM TCEP, and 1 mM $CaCl_2$.

Magnesium rescue of aspartate mutants' activity

Initial mutagenesis experiments with *At*PRORP1 indicated that all four fully-conserved aspartate residues near the active site are critical for activity (4). Given that our pH data was not consistent with PRORP using general base catalysis (1), we attempted to rescue the activity of the aspartate mutants with excess $MgCl_2$. For our new analysis, we used the low $MgCl_2$ conditions from the earlier paper or higher $MgCl_2$ (Fig. D-6) (4). At a 2 hour time point under both $MgCl_2$ conditions, WT enzyme cleaved the substrate to completion, while both D399A and D493A failed to generate measurable product. Interestingly, both D474A and D475A mutants catalyzed hydrolysis of substrate, with notable increases in product at higher concentrations.

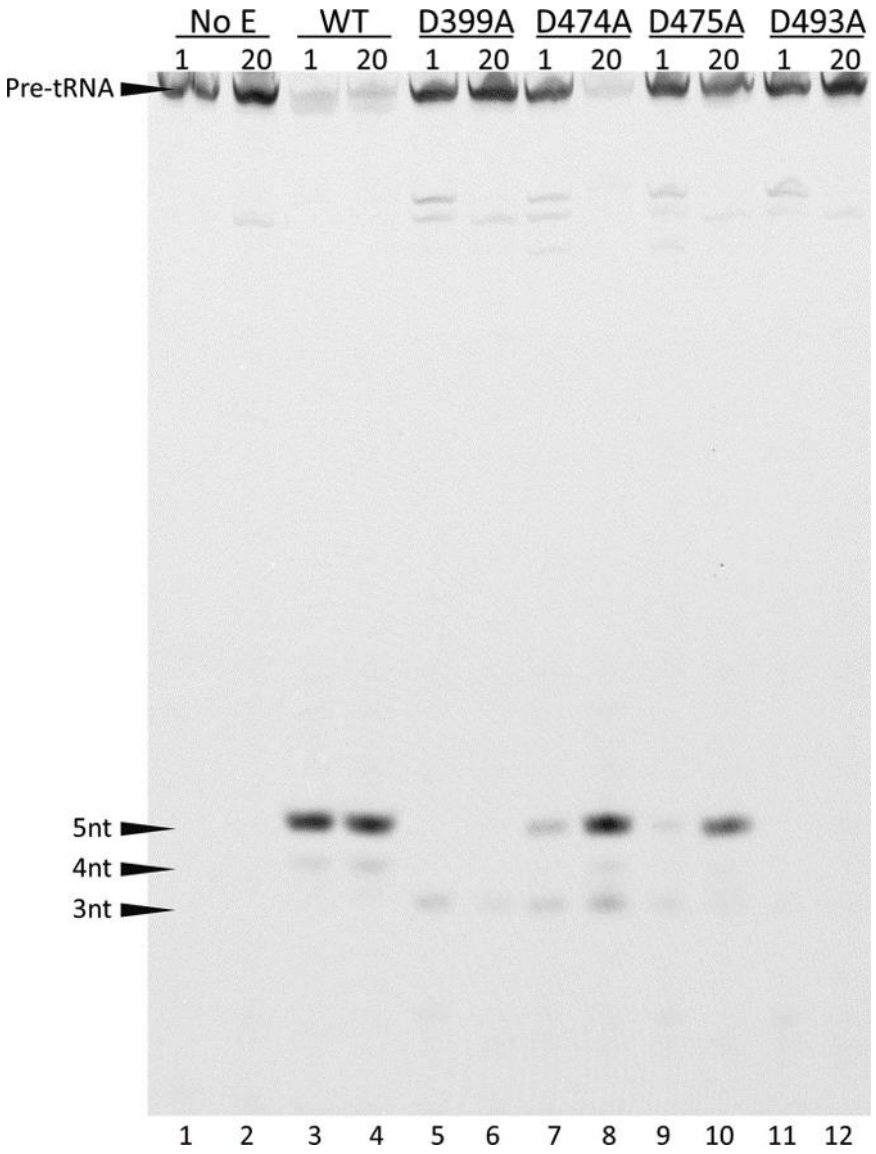


Figure D-6: Gel showing metal rescue of *At*PRORP1 aspartate mutants. *B*. *subtilis* pre-tRNA^{Asp} was incubated without enzyme (No E), with WT *At*PRORP1 (WT), or with one of the four aspartate mutants (D399A, D474A, D475A, D493A) for 2 hours with either 1 or 20 mM $MgCl_2$. The aspartate mutants had increased miscleavage behavior, as indicated by the shorter 4- and 3-nt products.

With the prior data in hand, we proceeded to quantify the metal rescue behavior. We measured the single-turnover k_{obs} at saturating WT, D474A, and D475A *At*PRORP1 and plotted these values versus the $MgCl_2$ concentration (Fig. D-7). Equation 4 was fit to the data. The mutants had > 9-fold increase in the Mg^{2+} $K_{1/2}$ and did not recover WT activity in the range of $MgCl_2$ concentrations we tested. However, with all three proteins we observe inhibition at very high $MgCl_2$ concentrations, so the relative recovery is difficult to assess. The data indicated that the primary role of Asp 474 and Asp 475 are to enhance the metal affinity at

the active site. We cannot yet conclude why D399A and D493A mutants could not be rescued, but we have proposed that these residues may be critical for active site structure. In the crystal structure of *At*PRORP1 (PDB 4g24), Asp 399 is interacting with two metal-bound waters, either of which might be the catalytic nucleophile, while Asp 493 is an inner sphere metal ligand and makes a hydrogen bond with the backbone amide nitrogen of Met 495.

$$k_{obs} = \frac{k_{max}^{Mg}[Mg^{2+}]^n}{[Mg^{2+}]^n + K_{1/2}^{Mg}} \quad (4)$$

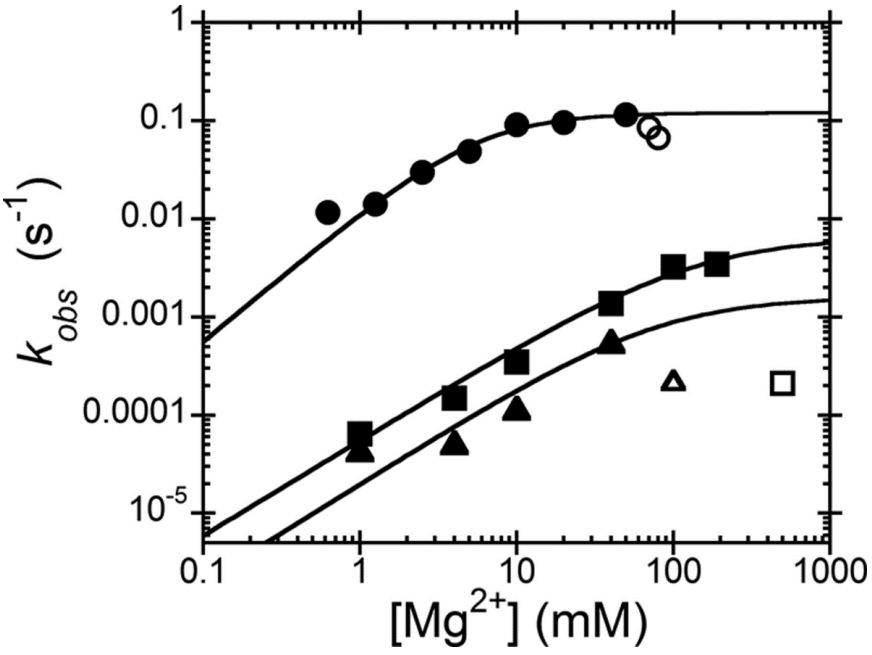


Figure D-7: Quantification of dependence of activity on the metal concentration for *At*PRORP1 aspartate mutants. Hyperbolic dependence on Mg^{2+} is observed for WT (circles), D474A (squares), and D475A (triangles) *At*PRORP1. Data indicating inhibition at high $MgCl_2$ concentrations (open symbols) were excluded for the purposes of fitting equation 4 to the data. The k_{max} values are 0.12 ± 0.01, ≥ 0.007, and ≥ 0.002 s^{-1} for WT, D474A, and D475A, respectively. The $K_{1/2}$ values are 10 ± 3.3 mM for WT and ≥ 90 mM for both D474A and D475A.

REFERENCES

1. Howard MJ, Klemm BP, Fierke CA. Mechanistic Studies Reveal Similar Catalytic Strategies for Phosphodiester Bond Hydrolysis by Protein-only and RNA-dependent Ribonuclease P. *J Biol Chem*. 2015;290(21):13454-64.
2. Good NE, Winget GD, Winter W, Connolly TN, Izawa S, Singh RM. Hydrogen ion buffers for biological research. *Biochemistry*. 1966;5(2):467-77.
3. Good NE, Izawa S. Hydrogen ion buffers. *Methods Enzymol*. 1972;24:53-68.
4. Howard MJ, Lim WH, Fierke CA, Koutmos M. Mitochondrial ribonuclease P structure provides insight into the evolution of catalytic strategies for precursor-tRNA 5 ' processing. *Proc Nat Acad Sci U S A*. 2012;109(40):16149-54.